

CMOS TELECOM DATA CONVERTERS

CMOS Telecom Data Converters

Edited by

Angel Rodríguez-Vázquez

*Institute of Microelectronics of Seville,
Spanish Microelectronics Center, IMSE-CNM (CSIC)
and Department of Electronics, University of Seville, Spain*

Fernando Medeiro

*Institute of Microelectronics of Seville,
Spanish Microelectronics Center, IMSE-CNM (CSIC)
and Department of Electronics, University of Seville, Spain*

and

Edmond Janssens

ST Microelectronics, Belgium



Springer Science+Business Media, LLC

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-1-4419-5382-7
DOI 10.1007/978-1-4757-3724-0

ISBN 978-1-4757-3724-0 (eBook)

Printed on acid-free paper

All Rights Reserved
© 2003 Springer Science+Business Media New York

Originally published by Kluwer Academic Publishers in 2003
Softcover reprint of the hardcover 1st edition 2003

No part of this work may be reproduced, stored in a retrieval system, or transmitted
in any form or by any means, electronic, mechanical, photocopying, microfilming, recording
or otherwise, without written permission from the Publisher, with the exception
of any material supplied specifically for the purpose of being entered
and executed on a computer system, for exclusive use by the purchaser of the work.

CONTENTS

Foreword	xix
Projects in the MSD Cluster	xxi
The book at a glance	xxiii
Contributors	xxvii
CHAPTER 1: NYQUIST-RATE CONVERTERS: AN OVERVIEW	1
<i>Roberto Rivoir</i>	
1. Introduction: Data Converters As Data Conversion Systems	1
2. Principles Of Data Conversion	4
2.1 Fundamental processes of analog-to-digital conversion	4
2.2 Quantization	5
2.3 Sampling, downsampling and oversampling	9
3. Architectures Of Nyquist Rate Converters	14
3.1 ADC classification	14
3.2 The integrating dual ramp ADC	15
3.3 The incremental ADC	17
3.4 Sigma-delta ADC versus Nyquist rate ADCs	19
3.5 The successive approximation ADC	21
3.6 The algorithmic ADC	24
3.7 The full flash ADC	25
3.8 The two-step flash ADC	27
3.9 Interleaving, pipelining	29
3.10 The pipeline ADC	30
3.11 The folding ADC	31
4. Future Directions And Conclusions	32
5. Acknowledgments	33

References	34
CHAPTER 2: SIGMA-DELTA CMOS ADCS: AN OVERVIEW OF THE STATE-OF-THE-ART	37
<i>Angel Rodriguez-Vázquez, Rocio del Rio, José M. de la Rosa, Ramón Tortosa, Fernando Medeiro and Belén Pérez-Verdú</i>	
1. Introduction	37
2. Architecture And Components Of $\Sigma\Delta$ Cs	42
3. Ingredients Of Sigma-Delta Modulators	45
3.1 Oversampling	46
3.2 Error Processing	48
3.3 Feedback	51
4. Classification Of $\Sigma\Delta$ M Ic Architectures	53
5. Discrete-Time Sigma-Delta Modulators	55
5.1 Single Quantizer Single-Bit Architectures	56
5.2 Single Quantizer Multi-Bit Architectures	60
5.3 $\Sigma\Delta$ M Architectures Containing Several Quantizers	64
5.4 BP DT $\Sigma\Delta$ Modulators	71
6. Continuous-Time Sigma-Delta Modulators	74
7. About The Influence Of Circuit Non-Idealities	79
8. Conclusions	81
9. References	81
9.1 General References	81
9.2 DT LP $\Sigma\Delta$ ICs	83
9.3 BP Sigma-Delta Modulator ICs	88
9.4 CT $\Sigma\Delta$ M ICs	89

CHAPTER 3: CURRENT-STEERING HIGH-SPEED D/A CONVERTERS FOR COMMUNICATIONS	93
<i>José Luis González and Eduard Alarcón</i>	
1. Introduction And State Of The Art	93
1.1 Architectural alternatives for Nyquist rate D/A converters	94
1.2 Current-steering D/A converter actual implementations	98
2. Error Sources In Current-Steering D/A Converters	101
2.1 Static and spectral performance metrics	101
2.2 Mismatch in the current cell	102
2.3 Finite output resistance	104
3. Circuit Implementation Of Current-Steering D/A Converters	107
3.1 The current cell	107
3.2 Dynamic errors	112
4. Architectures For Systematic Error Compensation	115
4.1 One dimensional switching sequences for systematic error compensation	115
4.2 Two dimensional switching sequences for systematic error compensation	117
4.3 Other systematic error compensation techniques	119
5. Optimum Segmentation In Current-Steering Segmented D/A Converters	121
5.1 Optimum segmentation based on area <i>vs</i> static performance	121
5.2 Optimum segmentation bases on spectral performance	125
6. Practical Design Of Current-Steering D/A Converters	129

6.1 Circuit sizing procedure	130
6.2 Layout issues	139
6.3 A high-speed high-accuracy current-steering D/A converter design example	143
7. References	145

CHAPTER 4: CMOS COMPARATORS **149**

Rafael Domínguez-Castro, Manuel Delgado-Restituto, Angel Rodríguez-Vázquez, José M. de la Rosa and Fernando Medeiro

1. Introduction	149
2. Overview Of Basic CMOS Voltage Comparator Architectures	150
2.1 Single-Step Voltage Comparators	151
2.2 Multi-Step Comparators	154
2.3 Regenerative Positive-Feedback Comparators	158
2.4 Pre-Amplified Regenerative Comparators	162
3. Architectural Speed-vs-Resolution Trade-Offs	164
3.1 Single-Step Comparators	164
3.2 Multi-Step Comparators	166
3.3 Regenerative Comparators	168
4. On The Impact Of The Offset	170
5. Offset-Compensated Comparators	174
5.1 Offset-Compensation Through Dynamic Biasing	174
5.2 Offset Compensation in Multi-Step Comparators	176
5.3 Residual Offset and Gain Degradation in Self-Biased Comparators	177
5.4 Transient Behavior and Dynamic Resolution in Self-Biased Comparators	179
6. Appendix I: Simplified Most Model	180
7. References	181

**CHAPTER 5: FOLDING/INTERPOLATING ADCs ANALOG PREPROCESSING
TECHNIQUES FOR HIGH-SPEED 8-BIT ADC** **183**
Koen Uyttendhove, J. Vandenbussche, G. Gielen and M. Steyaert

1. Introduction	183
2. Analog Preprocessing Techniques In ADCs	183
2.1 Folding techniques	184
2.2 Interpolation Techniques	186
2.3 Averaging Techniques	190
3. Architecture Of 8-Bit Converter	192
4. Systematic Design Of The A/D Converter	194
4.1 Design phase	194
5. Layout Of The Converter	205
6. Experimental Results	206
7. Conclusions	208
References	210

**CHAPTER 6: HIGH-SPEED FLASH ADCs - DESIGN ISSUES OF A 6 BIT, 1
GHz CMOS FLASH ADC** **213**
Koen Uyttendhove, Michiel Steyaert

1. Introduction	213
2. ADC Converter Architecture	214
2.1 High-level Offset Demand in Flash ADC	214
3. Building Blocks	215

3.1 Design of the Preamplifier	216
3.2 Design of Comparator	217
3.3 Design of the Resistor Ladder	221
3.4 Design of the Digital Logic	225
4. Experimental Results	233
5. Conclusion	236
References	239

CHAPTER 7: LOGARITHMIC ANALOGUE-TO-DIGITAL CONVERTERS 241*Jorge Guilherme and Joao Vital*

1. Introduction	241
2. Non-Uniform Quantization Methods	242
2.1 Low speed non-uniform quantizers	243
2.2 High-speed non-uniform quantizers	245
3. Transfer Characteristic	246
4. Signal-To-Noise Ratio	248
5. Pipeline Logarithmic ADC	250
5.1 Logarithmic transformation	250
5.2 Logarithmic pipeline ADC	254
5.3 Symmetrical pipeline logarithmic ADC	255
5.4 Architecture improvements	257
6. Linearity Characterization	259
7. Digital Calibration	261
8. Two-Step Flash Pipeline ADC	264
9. A Logarithmic Pipeline ADC With 80 db Dynamic Range	270
10. Summary	272

11. References	273
CHAPTER 8: SINGLE-LOOP MULTI-BIT SIGMA-DELTA MODULATORS	277
<i>Yves Geerts, Michiel Steyaert, Willy Sansen</i>	
1. Introduction	277
2. Architectures	278
2.1 Multi-Bit Single-Loop Topologies	278
2.2 Performance Comparison	281
2.3 Linearity Issues of Multi-Bit $\Delta\Sigma$ Converters	283
3. Implementation Of The Integrator	285
4. Circuit Non-Idealities	288
4.1 Finite OTA gain	289
4.2 Finite dominant closed-loop pole of the OTA	292
4.3 Resistance of the switches	293
4.4 Slew-rate effects	297
4.5 Offset effects in the quantizer	298
5. Design Example	299
6. Conclusion	303
References	303
CHAPTER 9: HIGH-ORDER CASCADE MULTI-BIT $\Sigma\Delta$ MODULATORS	307
<i>F. Medeiro, R. del Río, J.M. de la Rosa, B. Pérez-Verdú and A. Rodríguez-Vázquez</i>	
1. Introduction	307
2. Expandible, Modular Cascade Multi-Bit Modulators	310

2.1 Cascade multi-bit $\Sigma\Delta$ modulators	310
2.2 Architecture optimization	312
3. Technology-driven scenario for cascade multi-bit sigma-delta modulators	316
3.1 Non-ideal operation	316
3.2 Impact of deep-submicron features	320
3.3 Comparison among cascade architectures	329
4. Design Example	333
4.1 Switched-capacitor implementation	333
4.2 Building block specifications	335
4.3 Building block implementation	337
4.4 Results	340
References	341
CHAPTER 10: CONTINUOUS-TIME SIGMA-DELTA FOR IF	345
<i>F. Gerfers, Y. Manoli, M. Ortmanns</i>	
1. Introduction	345
2. DT/CT Modulator Equivalence	346
2.1 Modified Z-Transformation	346
2.2 Equivalent loopfilters principle: the impulse-invariant transformation	348
2.3 Extension for Multi Loop Modulators	349
3. Influence Of Circuit Non-Idealities	352
3.1 Integrator Non-Idealities	352
3.2 Temporal Non-Idealities	357
3.3 Other Non-Idealities	360
4. Error Compensation And Cancellation	361

4.1 Cancellation of Gain Mismatch in Cascaded Modulators	361
4.2 CT Modulators with reduced Clock Jitter Sensitivity	362
5. Implementation Of Low Power CT $\Sigma\Delta$ Modulators	367
5.1 Figure of Merit, FOM	367
5.2 Design Example: A 10bit-25kHz CT $\Sigma\Delta$ Modulator	370
5.3 Low-Power Third Order $\Sigma\Delta$ Modulator	372
5.4 Measurement Results	372
6. Conclusion	373
References	374

**CHAPTER 11: BANDPASS SIGMA-DELTA A/D CONVERTERS:
FUNDAMENTALS, ARCHITECTURES AND CIRCUITS** **379**

*José M. de la Rosa, Belén Pérez-Verdú, Rocío del Río, Fernando Medeiro and
Angel Rodríguez-Vázquez*

1. Introduction	379
2. The Digital Wireless Communications Universe	380
2.1 The ideal digital wireless transceiver	381
2.2 Overview of wireless digital radio receiver architectures	381
2.3 IF A/D Conversion – Bandpass $\Sigma\Delta$ modulators	384
3. Basic Concepts Of Bandpass $\Sigma\Delta$ A/D Converters	384
3.1 Signal passband location	387
3.2 Decimation for bandpass $\Sigma\Delta$ ADCs	388
4. Synthesis Of Bandpass $\Sigma\Delta$ Modulator Architectures	390

4.1 The lowpass-to-bandpass transformation method: LP-to-BP	390
4.2 Optimized synthesis of $N_{TF}(z)$	393
4.3 Quadrature bandpass $\Sigma\Delta$ modulators	394
4.4 N -path bandpass $\Sigma\Delta$ modulators	397
4.5 Synthesis of continuous-time bandpass $\Sigma\Delta$ modulators	397
5. Building Blocks And Error Mechanisms In BP $\Sigma\Delta$ Ms	399
5.1 From integrators to resonators	399
5.2 Quantization noise shaping degradation due to circuit errors	402
5.3 Harmonic distortion in bandpass $\Sigma\Delta$ modulators	406
5.4 Thermal noise	409
5.5 Jitter noise	410
5.6 Excess loop delay in continuous-time BP $\Sigma\Delta$ Ms	412
6. State-Of-The-Art Bandpass $\Sigma\Delta$ ADCs	412
7. References	416

**CHAPTER 12: DECIMATION FILTER DESIGN FOR SIGMA-DELTA
CONVERTERS** **421***Valentino Liberali, Roberto Rossi and Guido Torelli*

1. Introduction	421
2. Basic Concepts Of Decimation	422
2.1 Decimation in the z domain	422
2.2 The commutative rule	427
3. Design Flow	428
4. Architecture	429

4.1 Multi-stage vs. single-stage	431
4.2 First stage	434
4.3 Equalizer	437
4.4 Band filter	438
5. Design Entry	439
5.1 Hierarchy	439
5.2 Granularity	439
5.3 Clocking strategy	440
6. Synthesis	440
6.1 Constraints	441
6.2 Synthesis strategies	441
6.3 Gate-level simulations	442
6.4 Layout	442
7. A Design Example	442
7.1 Top level	442
7.2 Sinc filter	445
7.3 Equalizer	447
7.4 Band filter	447
8. References	447
CHAPTER 13: LINE DRIVERS : EFFICIENCY, LINEARITY, BANDWIDTH	449
<i>Tim Piessens and Michiel Steyaert</i>	
1. Introduction	449
2. Wire-Line Communication	451
2.1 xDSL-Technologies	451
2.2 Basic Properties of Multi-Tone Channels	453
3. Linear Power Amplifiers	456

Appendix 14.III Dynamic Range	515
5. References	517

**CHAPTER 15: DESIGN METHODOLOGIES FOR SIGMA-DELTA
CONVERTERS** 523

*Francisco V. Fernández, Rocío del Río, Rafael Castro-López, Oscar Guerra,
Fernando Medeiro and Belén Pérez-Verdú*

1. Introduction	523
2. Top-Down Design Methodology	524
3. Simulation	525
3.1 Analytical models for event-driven behavioral simulation	528
3.2 Look-up tables	539
3.3 Postprocessing	540
3.4 Implementation and examples	541
4. Top-Down Synthesis	543
4.1 Synthesis at the modulator level	544
4.2 Synthesis at the cell level	550
5. Bottom-Up Layout Construction And Verification	553
6. References	557

**CHAPTER 16: SYSTEMATIC COMPUTER – AIDED DESIGN METHODOLOGY
FOR CMOS DATA CONVERTERS** 561

Georges Gielen and Kenneth Francken

1. Introduction	561
2. Systematic Design Flow For Data Converters	563
3. Generic Behavioral Modeling For The Top-Down Phase	565
4. Sizing Synthesis Of The Data Converter	566

4.1 Architectural-level synthesis of a $\Delta\Sigma$ A/D converter	568
4.2 Systematic design of a Nyquist-rate D/A converter	575
5. Extracted Behavioral Model For Bottom-Up Verification	583
6. Conclusions	585
Acknowledgements	586
References	586

Foreword

This book is the second in a series of three dedicated to advanced topics in Mixed-Signal IC design. It is one of the results achieved by the Mixed-Signal Design Cluster, an initiative launched in 1998 as part of the TARDIS project, funded by the European Commission within the ESPRIT-IV Framework. This initiative aims to promote the development of new design and test methodologies for Mixed-Signal ICs, and to accelerate their adoption by industrial users.

As Microelectronics evolves, Mixed-Signal techniques are gaining importance due to the wide spread of applications where an analog front-end is needed to interface a complex digital-processing subsystem. In this sense, Analog and Mixed-Signal circuits are recognized as a bottleneck for the market acceptance of Systems-On-Chip, because of the inherent difficulties involved in the design and test of these circuits. Specially, problems arising from the need to have high-performance interfaces between analog and digital components are a limiting factor for the performance achievable in mixed-signal designs for System-On-Chip.

The Mixed-Signal Cluster has been formed by a group of 11 Research and Development projects, plus a specific action to promote the dissemination of design methodologies, techniques and supporting tools developed within the Cluster projects. The whole action, ending in July 2002, has been assigned an overall budget of more than 8 million EURO.

The novelty of the TARDIS initiative is that in addition to the standard R&D work, the participating projects have a compromise to publicize the new methodological results obtained in the course of their work. A Cluster Coordinator, Instituto de Microelectrónica de Sevilla, in Sevilla (Spain) has the role to coordinate and promote actions to carry out effectively the dissemination work and foster cooperation between the participating projects. All public results from the dissemination action are available from the Cluster Web site (<http://www.imse.cnm.es/esd-msd>).

Activities of projects in the Cluster have been focused on four main areas (Substrate Noise Coupling, Advanced Data Converters, Testability and Special Technologies). This book addresses the design of Data Converters and incorporates the results achieved by the Cluster projects, complemented by contributions from external experts who have participated in activities organized by the Cluster.

We hope that readers will find this book useful, and we would like to thank all partners of the MSD Cluster for contributing to the success of the initiative. Special thanks are given to all the authors and to the editors for their effort.

José Luis Huertas, Juan Ramos-Martos

Sevilla, September 2002

Projects in the MSD Cluster

- **ABACUS: Active Bus Adaptor and Controller for Remote Units.** The objective of this project is the development of an integrated circuit for space applications, that implements the analog/digital interface between the spacecraft On-Board Data Handling (OBDH) bus, and the Remote Terminal Units (RTUs). The design will use $0.8\mu\text{m}$ SOI technology.
- **BANDIT: Embedding Analog-to-Digital Converters on Digital Telecom ASICs.** The goal of BANDIT is to develop a general design methodology for embedding high-speed analog/digital converters (ADCs) on large digital telecom ASICs, with special attention to the problems caused by mixed-signal integration.
- **HIPADS: High-Performance Deep Sub-micron CMOS Analog-to-Digital Converters using Low-Noise Logic.** The aim of this project is to develop different A/D Converters in deep sub-micron digital CMOS process, using a new Current Steering Logic (CSL) family approach that has the property of inducing a very low substrate noise. The converters are intended to become integrated components of larger systems, and should be considered presently as products under specifications covering end-user applications.
- **MADBRIC: Mixed Analog-Digital Broadband IC for Internet Power-Line Data Synchronous Link.** The project main objective is the development of prototype building blocks of a chipset for high-speed communications through the power lines, that will improve achievable data rates using state of the art mixed-signal integrated circuits and DSP techniques.
- **MIXMODEST: Mixed Mode in Deep Submicron Technology.** The technical target of the MIXMODEST project is to develop design techniques that permit the implementation of mixed-signals systems in the most advanced $0.35\mu\text{m}$ and $0.25\mu\text{m}$ deep sub-micron digital CMOS technology.
- **OPTIMISTIC: Optimisation Methodologies in Mixed-Signal Testing of ICs.** The OPTIMISTIC project, concerned with Optimisation Methodologies in Mixed-Signal Testing of ICs, aims at the development and introduction of advanced test generation in mixed-signal IC design. Building upon existing advanced tools for control and test systems, a new approach is to be developed that will allow the mixed-signal chip designer to take large responsibility in the generation of test as part of the design activity.
- **RAPID: Retargetability for Reusability of Application-Driven Quadrature D/A Interface Block Design.** This project is concerned with the development of an advanced methodology for the design of a mixed-signal application-driven quadrature D/A interface sub-system, aiming at its reusability by a retargeting procedure with minimal changes to their structural sub-blocks.
- **SUBSAFE: Substrate Current Safe Smart Power IC Design Methodology.** The overall technical objective of this project is to develop a design methodology that employs device and circuit simulation to assure IC digital functionality under current injection in the substrate produced by forward bias conditions in N-wells (i.e. during switching of power stages driving inductive loads). The design methodology will change from the current largely empirical approach to Computer-Aided Design guided critical parameter evaluation, validated by a relatively small number of measurements.
- **SYSCONV: Systematic Top-Down Design and System Modeling of Oversampling Converters.** This project develops a system-level model for oversampling delta-sigma converters suitable for use in mixed-signal system simulations and verifications. It addresses the development of a model of the entire converter as a block on its own, that can then be used in efficient mixed-signal system simulations where the converter is only a

block in the overall system.

- **TERMIS: High-Temperature/High-Voltage Mixed Signal SOI ASICs for Aerospace Applications.** The project addresses the development of a fully integrated high-voltage driver IC for two different electromagnetic micro-motors which are dedicated for satellite applications. Each circuit, in die form, will be packaged in the corresponding micro-motor. The systems must operate at 200°C under a 30V power supply and must survive space irradiation.
- **VDP: Video Decoder Platform.** This project develops a prototype video decoder platform. The result will be an IC that captures video signals and decodes the information for use in, for instance digital TV, set top boxes, and PC video capture. It will exploit innovative architectures trading signal to noise ratio versus accuracy, decoding both analog and digital video sources.

The book at a glance

Surprisingly quickly, in only a few decades, telecommunications in the modern sense of the word have moved from being a mere fashionable concept to an irreplaceable tool. One only needs to imagine one's everyday life without the ability to exchange information from a distance in order to realize the importance of those gadgets that go unnoticed in the daily routine. In spite of this revealing fact, our most recent history tells us to humbly admit that probably telecom technology is just in its earliest days. Just like radio society was revolutionized by TV, and conventional telephony by mobile phones and internet, so will new devices appear which will revolutionize communications. Furthermore, forecasts are that telecom will be embedded with many-colored systems to provide us with an "intelligent" environment ready to interact.

Hopefully, we will not have to wait too long. As technology advances, the time needed for a standard to reach maturity and decline is getting ever shorter. For instance, the internet has required a bit more than a decade to move from military to scientific usage and less than that to become one of the most popular media. Fifteen years ago a mobile phone was an exclusive, briefcase-sized device, which required a car battery to operate.

Undoubtedly, Microelectronics has been crucial in speeding up these changes. During the last quarter of a century, global and personal media have been inspired and supported by the parallel evolution of digital integrated systems and their fabrication processes. Although digital signal processing has opened doors (and continues doing so) to new communication standards, improving speed, reliability, range, and portability, we cannot forget that the act of communication itself is fundamentally analog. Sometimes it is the channel which requires analog signals to be transmitted, at other times it is the human being who naturally handles these signals. Thus, interfacing the analog and digital worlds should be considered of fundamental importance for present and future telecom devices. In fact, modern successful telecom systems exhibit a smart balance between complex digital processing and high-performance analog functions; yet, it is well-known that, as fabrication processes evolve, this balance tips in digital's favour.

As the trend to system-on-chip solutions gradually consolidates, including an ever-increasing usage of digital signal processing, a concurrent enhancement of the analog front-end performance is required. Moreover, such enhanced performances must be achieved under the constraints imposed by digitally driven technology roadmaps, with low-voltage supplies, poor-performance (and often badly characterized) devices, ... and above all the unavoidable presence of noisy digital circuits. It is under the pressure of these challenges that analog designers must

sharpen their wits to overcome the problems. The results of recent research and development projects and hundreds of publications in journals and conferences confirm that solutions can be found.

Both the designers' ingenuity and slow, but precious, improvements of digital CMOS technologies keep mixed signal design in continuous evolution. Maybe there is no better reason than this for yet another book on CMOS telecommunication converters. This is one of the purposes of this book, namely compiling the freshest design activities of a group of highly skilled designers; introducing powerful analog-to-digital converter architectures and robust design techniques from a sometimes controversial but definitely complementary point of view. This is the case of Chapters 5 to 11, aimed at presenting the latest research achievements on four converter architectures which play a leading role in modern communication systems, namely:

- Oversampling Sigma-Delta (or delta-sigma),
- Pipeline
- Folding/Interpolating
- Flash

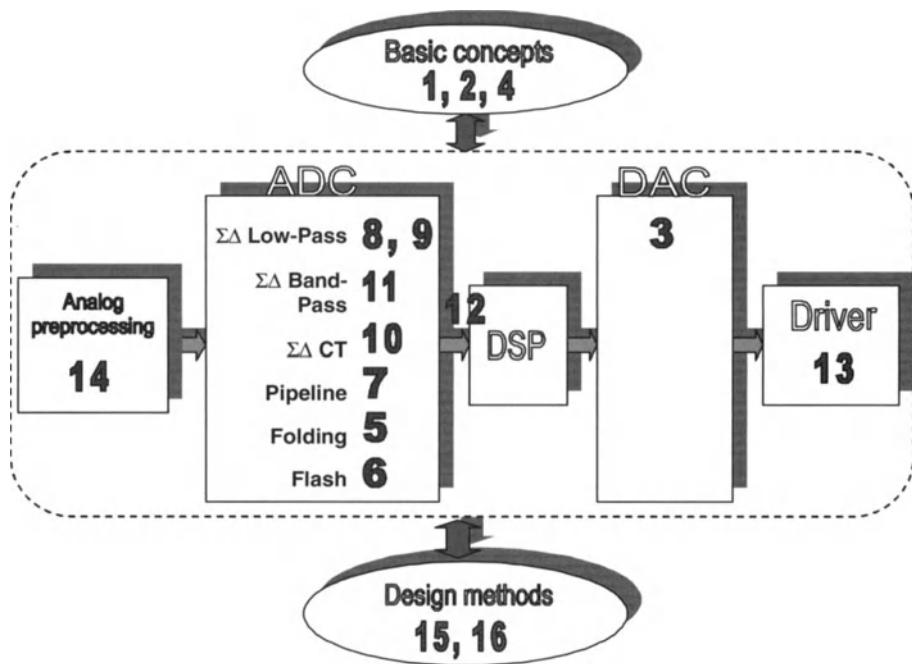
However, in the highly specialized professional life that we all lead, it is more and more difficult to have the expertise required for taking the maximum advantage of the latest technical achievements. In order to prevent novices from getting frustrated with pages of involved equations and condensed, high-level explanations, the authors have been asked to write their chapters in a structured and comprehensive way. Furthermore, specific chapters devoted to presenting basic concepts and overviews are also included. This is the case of Chapter 1, for Nyquist converters, Chapter 2, for oversampling sigma-delta converters, and Chapter 4, for CMOS comparators (the very analog building block of analog-to-digital converters).

Besides analog-to-digital converters, other embedded functions are needed to realize modern communication systems. Some of these functions are also addressed in this book, namely: digital-to-analog converters (Chapter 3), analog preprocessing (Chapter 14), digital post-processing (Chapter 12) and line driving (Chapter 13).

Last but not least, design methods are of prime importance for quick, efficient implementation of any integrated circuit and specifically for the functions covered in this book. This is the topic of Chapters 15 and 16.

To sum things up, the figure below provides a view of the topics covered in this book, where the chapter numbers have been placed on a simplified, conceptual block diagram of a communication system, reflecting the main topics covered in each of them.

Depending on reader interest and background, different reading sequences are



suggested. Those with a general background in analog circuit design, but limited expertise on data converters may start with Chapters 1, 2 and 4, continue with Chapters 15 and 16, and then select from the rest of chapters those that best fit their specific interests or needs.

Another sequence is suggested for readers specifically interested in sigma-delta converters, which, owing to their increasing relevance for both broadband wireline and narrowband wireless communications, deserve the attention of several chapters. For these readers we suggest starting with Chapters 2 and 4, continuing with Chapters 8 and 9, then 11, and finally 10. Such a reading sequence should be completed with the specific design methods in Chapter 15, and the digital post-processing included in Chapter 12.

However, readers interested in Nyquist converters should follow this path: Chapters 1, 3 and 4, then Chapters 6, 7 and 5, and finally Chapter 16.

Although it is clear to us that a book on analog design will never be 100% satisfactory, we hope that your feeling as a reader will not reach the other digital state (0%), but something in between.

Ángel Rodríguez-Vázquez

Fernando Medeiro

Edmond Janssens

Sevilla, January 2003

CMOS Telecom Data Converters

Eduard Alarcón

Department of Electronic Engineering
ETS d'Enginyeria de Telecommunicació de Barcelona
Universitat Politècnica de Catalunya - Nonlinear Circuits and Systems Group
Address: Office.105, Building C4, c/ Gran Capità s/n,
08034 Barcelona SPAIN
Phone: (34) (93) 401 56 78
Fax: (34) (93) 401 67 56
e-mail: ealarcon@eel.upc.es

Rafael Castro-López

Institute of Microelectronics of Seville-CNM.IMESE
Address: Avda. Reina Mercedes s/n Edif. CICA
41012 Seville (SPAIN)
Phone: + 34 955 056 666
Fax: + 34 955 056 686
e-mail: Rafael.Castro@imse.cnm.es

José M. de la Rosa

Institute of Microelectronics of Seville-CNM.IMESE
Address: Avda. Reina Mercedes s/n Edif. CICA
41012 Seville (SPAIN)
Phone: + 34 955 056 666
Fax: + 34 955 056 686
e-mail: Jose.de.la.Rosa@imse.cnm.es

Rocío del Río

Institute of Microelectronics of Seville-CNM.IMESE
Address: Avda. Reina Mercedes s/n Edif. CICA
41012 Seville (SPAIN)
Phone: + 34 955 056 666
Fax: + 34 955 056 686
e-mail: Rocio.del.Rio@imse.cnm.es

Manuel Delgado-Restituto

Institute of Microelectronics of Seville-CNM.IMESE
Address: Avda. Reina Mercedes s/n Edif. CICA
41012 Seville (SPAIN)
Phone: + 34 955 056 666
Fax: + 34 955 056 686
e-mail: **Manuel.Delgado@imse.cnm.es**

Rafael Domínguez-Castro

Institute of Microelectronics of Seville-CNM.IMESE
Address: Avda. Reina Mercedes s/n Edif. CICA
41012 Seville (SPAIN)
Phone: + 34 955 056 666
Fax: + 34 955 056 686
e-mail: **Rafael.Dominguez-Castro@imse.cnm.es**

Francisco V. Fernández

Institute of Microelectronics of Seville-CNM.IMESE
Address: Avda. Reina Mercedes s/n Edif. CICA
41012 Seville (SPAIN)
Phone: + 34 955 056 666
Fax: + 34 955 056 686
e-mail: **Francisco.Fernandez@imse.cnm.es**

Kenneth Francken

Katholieke Universiteit Leuven, ESAT-MICAS
Dept. Elektrotechniek
Address: Kasteelpark Arenberg 10, Room 91.21
B-3001 LEUVEN, Belgium
Phone: +32 16 321076
Fax : +32 16 321975
e-mail : **Kenneth.Francken@esat.kuleuven.ac.be**

Yves Geerts

ST Microelectronics
e-mail: **yves.geerts@st.com**

Friedel Gerfers

University of Freiburg

Georges Gielen

Katholieke Universiteit Leuven, ESAT-MICAS

Dept. Elektrotechniek

Address: Kasteelpark Arenberg 10, Room 01.38

B-3001 LEUVEN, Belgium

Phone: +32 16 321047

Fax : +32 16 321975

e-mail : Georges.Gielen@esat.kuleuven.ac.be

José Luis González

UPC Dept. Electronic Engineering

Address: Jordi Girona 1-3 C4 Building - Campus Nord Barcelona

08034 Barcelona - Spain

Phone: +34 934016748

Fax: +34 934016756

e-mail: jlgonzalez@eel.upc.es

Oscar Guerra

Institute of Microelectronics of Seville-CNM.IMESE

Address: Avda. Reina Mercedes s/n Edif. CICA

41012 Seville (SPAIN)

Phone: + 34 955 056 666

Fax: + 34 955 056 686

e-mail: Oscar.Guerra@imse.cnm.es

Jorge Guilherme

Escola Superior Tecnologia Tomar

Departamento de Engenharia Electrotecnica Industrial

Address: Quinta do Contador, Estrada da Serra

2300 Tomar, Portugal

Phone: +351 49 328150

Fax: +351 49 328197

e-mail: guilherm@chipidea.com

Valentino Liberali

University degli Studi di Milano
Dipartimento di Tecnologie dell'Informazione (DTI)
Address: Via Bramante 65
26013 CREMA (CR) - Italy
Phone: +39-02.503.30071 or +39-0373.898.071
Fax: +39-02.503.30010 or +39-0373.898.01
e-mail: liberali@dti.unimi.it

Yiannis Manoli

Albert-Ludwig-Univ.-IMTEK
Address: Georges-Koehler-Allee 102
D-79110 Freiburg, Germany
Phone.: + 49.761.203.7590
Fax: +49.761.203.7592
e-mail: manoli@imtek.uni-freiburg.de

Fernando Medeiro

Institute of Microelectronics of Seville-CNM.IMESE
Address: Avda. Reina Mercedes s/n Edif. CICA
41012 Seville (SPAIN)
Phone: + 34 955 056 666
Fax: + 34 955 056 686
e-mail: Fernando.Medeiro@imse.cnm.es

Maurits Ortmanns

Albert-Ludwigs-University
Address: Georges-Koehler-Allee 102
D-79110 Freiburg, Germany
Phone: +49.761.203.7596
Fax: +49.761.203.7592
e-mail: maurits@imtek.uni-freiburg.de

Belen Pérez-Verdú

Institute of Microelectronics of Seville-CNM. IMSE
Address: Avda. Reina Mercedes s/n Edif. CICA
41012 Seville (SPAIN)
Phone: + 34 955 056 666
Fax: + 34 955 056 686
e-mail: Belen.Perez@imse.cnm.es

Tim Piessens

Katholieke Universiteit Leuven - ESAT-MICAS
Dept. Elektrotechniek
Address: Kasteelpark Arenberg 10, Room 91.18
B-3001 LEUVEN, Belgium
Phone: +32 16 321149
Fax : +32 16 321975
e-mail : Tim.Piessens@esat.kuleuven.ac.be

Roberto Rivoir

Atmel Rousset
Wireless Design Group Zone Industrielle
Address: Rousset Cedex 1
3106 FRANCE
Phone: 0033.442536343
Fax: 0033.442536001
e-mail: rrivoir@atmel.fr

Angel Rodríguez-Vázquez

Unit Head High-Performance Mixed-Signal VLSI Ics
Institute of Microelectronics of Seville-CNM. IMSE
Address: Avda. Reina Mercedes s/n Edif. CICA
41012 Seville (SPAIN)
Phone: + 34 955 056 666
Fax: + 34 955 056 686
e-mail: Angel.Rodriguez@imse.cnm.es

Roberto Rossi

Department of Electronics
University of Pavia
Address: Via Ferrata, 1
27100 Pavia, Italy
e-mail: roberto@ele.unipv.it

Willy Sansen

Katholieke Universiteit Leuven, ESAT-MICAS
Dept.Elekrotechniek
Address: Kasteelpark Arenberg 10, Room 01.22
B-3001 LEUVEN, Belgium
Phone : + 32 16 321077
Fax : + 32 16 321975
e-mail : Willy.Sansen@esat.kuleuven.ac.be

Michiel Steyaert

Katholieke Universiteit Leuven, ESAT-MICAS
Dept. Elektrotechniek
Address: Kasteelpark Arenberg 10, Room 01.21
B-3001 LEUVEN, Belgium
Phone: +32 16 321059
Fax : +32 16 321975
e-mail : Michiel.Steyaert@esat.kuleuven.ac.be

Guido Torelli

Department of Electronics
University of Pavia
Address: Via Ferrata, 1
27100 Pavia, Italy
Phone: + 39 0382505215
e-mail: guido@ele.unipv.it

Ramón Tortosa

Institute of Microelectronics of Seville-CNM. IMSE
Address: Avda. Reina Mercedes s/n Edif. CICA
41012 Seville (SPAIN)
Phone: + 34 955 056 666
Fax: + 34 955 056 686
e-mail: **Ramon.Tortosa@imse.cnm.es**

Koen Uyttenhove

ESAT-MICAS Katholieke Universiteit Leuven
Phone: +32-16-321070
Fax: + 32-1632 1975
e-mail: **uyttenho@esat.kuleuven.ac**

Jan Vandenbussche

Katholieke Universiteit Leuven, ESAT-MICAS
e-mail: **jan.vandenbussche@ansem.com**

Joao Vital

Chipidea – Microelectrónica, S.A.
Address: Taguspark - Ed. Inovacao IV, Sala 733
2780-920 Porto Salvo, Portugal
Phone: + 351-210336302
Fax:+ 351-210336396
e-mail: **jvital@chipidea.com**

Chapter 1

Nyquist-rate Converters: An Overview

Roberto Rivoir

1.1 INTRODUCTION: DATA CONVERTERS AS DATA CONVERSION SYSTEMS

Digital-to-analog and analog-to-digital converters are among the most important technology enablers for integrated circuits for communication applications. From a theoretical point of view, and in order to better study the different architectures, it is convenient to consider data converters as stand-alone IP (Intellectual Property) building blocks, and classify them according to the most important principles of operation. This leads us, for the sake of simplicity, and as it has been done in the current textbook, to define two main categories: Nyquist-rate and oversampled converters. This extreme simplification needs, however, some clarifications, that will be carried out in the paragraph dedicated to the converters architectures. Moreover, the “stand-alone IP-block” assumption should not lead to misleading interpretations, since in practical applications it is highly improbable to find spare ADCs or DACs used as stand-alone cells, to implement self-consistent functions. On the contrary, and especially in Systems-On-Chips (SOCs), ADCs and DACs are the core elements of much more highly integrated functions where, around the key operation of converting data from the analog (digital) domain to the digital (analog) domain, a number of other vital operations are implemented as well: i) integrate those building blocks that are indispensable for the converter to work properly (*bandgap references, current references, reference buffers*), ii) interface it to the external world (*input buffers, programmable gain amplifiers, output buffers, power amplifiers, line drivers*), iii) ensure proper conditions to reach a target performance (*dc-dc converters, high PSRR voltage regulators, low jitter phase-locked loops*), iv) provide a number of additional features to complete the functionality of the overall data conversion channel (*antialiasing, reconstruction filtering, analog and/or digital channel filtering, serial/parallel I/O ports, offset compensation, overflow detection, power-up-down management, thermal protection, ...*), and finally, sometimes neglected but extremely important, v) add that hardware overhead which allows a complete and cost effective testability of the system (*SCAN test, BIST analog,...*). Given a specific application, which determines consequently a

specification for the whole data conversion system in terms of static and dynamic parameters (power supply, current consumption, power supply rejection ratio, integral and differential non-linearity, signal-to-noise ratio, total-harmonic distortion,...), the converter of choice will fall within a restricted number of architectures: oversampled or Nyquist rate, and in the last case slow speed, very high resolution (*dual ramp, incremental*), or medium speed, n-clock cycles (*algorithmic, successive approximation*), or finally very high speed, 1-2 clock cycles converters (*pipeline, two-step flash, full flash, folding*). Usually the data converter represents the most difficult building block to be developed, and the key knowledge of the designer concerning one of its possible architectures, determines not only the choice of the converter itself, but also the development of all remaining auxiliary functions around it. The specification of the data conversion channel directly translates into the one of the converter, and the auxiliary blocks have to be properly sized, in order to guarantee and preserve the full performance. The example of Fig. 1.1 can help to gain a more in depth understanding of the subject.

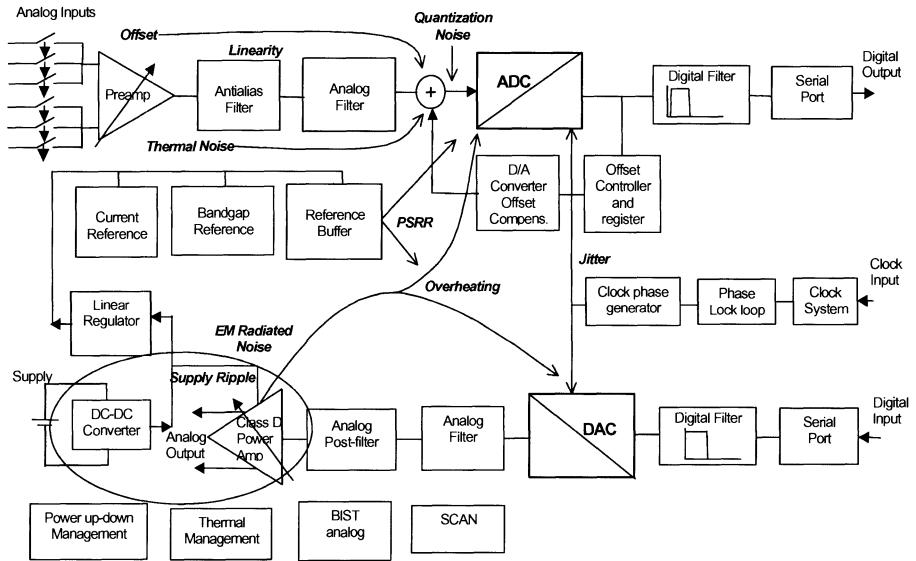


Figure 1.1: Audio codec as example of ADC and DAC data conversion system.

An audio codec is represented in Fig. 1.1, where the core ADC and DAC IP blocks are integrated with a number of other blocks to implement the whole function. In this kind of circuit, typically fabricated with inexpensive CMOS digital technology, data converters are based on sigma-delta (oversampled) [1,2] or successive approximation (Nyquist-rate) [3] architectures. As said before, the converter type determines the architecture of the overall conversion channel. In the case of sigma-delta A/D converters, a simple and low-order filtering before conversion, featuring just the anti-aliasing requirement, will be performed, while the required voice-band filtering, after decreasing the sample rate with multi-rate decimation filters, will take place in the digital domain (similar considerations

apply for the DAC channel). In the case of successive approximation ADC, a full voice-band filtering can be performed in the analog domain, for example through switched-capacitor analog filters. It should be stressed that these architectural choices do not strictly depend only on the data converter type, since they are also a consequence of the technology. Sigma-delta converters are often used in CMOS deep sub-micron, digital dedicated technologies (0.18um or lower feature size), often with no analog process options (double-poly capacitors), or without an available analog characterization. Since these technologies offer very high gate density, it is convenient to use both: a robust and technology insensitive data conversion technique like sigma-delta, in conjunction with low area occupation digital filtering. On the other hand, for a larger feature size, but more mature and well characterized technology, a voice-band filtering in the analog domain, in conjunction with Nyquist-rate conversion, can be the most convenient choice.

As Fig. 1.1 highlights, there are several interactions among different circuits, which determine the overall performance of the system. Assuming that a dynamic performance of SNR=70dB is targeted, the ADC converter will be the main block, but not the only one, affecting this performance. For high level input signals, it can be expected that the quantization noise of the ADC will be the dominant term, but as the input signal will be weak, and the programmable gain amplifier will be set to high gain, the equivalent thermal noise at the ADC input will be not negligible, and the SNR will worsen consequently. Still in this situation, the offset voltage of the microphone amplifier can be also very high, and could prevent the ADC from correctly reaching its peak SNR due to an overload problem (signal exceeding the full scale of the converter). In order to exploit the full dynamic range of the converter, an offset compensation circuit may be necessary. Jitter associated with the master clock, coming into the most sensitive parts of data converters, is another potential cause of performance reduction: it may worsen the SNR and increase the idle noise. Coming to the DAC channel, there is the tendency to integrate on-chip high output power amplifiers, capable of driving low impedance speakers. In case the output amplifier will be a class AB one, we should remember that the practical efficiency of this amplifier will be around 50%, so that nearly half of the power will be dissipated on-chip, implying potential overheating and thermal gradients, which can be responsible of performance degradations. If a class D amplifier is used, then higher efficiency will imply less overheating on-chip, but an increase of both electromagnetic radiated noise, and power supply noise. These last two phenomena can be critical if the right countermeasures are not taken, like filtering out of the radiated noise, and defining a proper power supply management scheme.

As a conclusion of this Introduction, the reader should retain the basic idea that for a particular need of data conversion, the choice of the converter type will strongly influence the architecture of the overall system. Since the design of the converter cannot be separated from the design of the full system, this implies, from the beginning of the project, the involvement of a design team with multidisciplinary expertise on analog, digital, mixed-mode, design-for-testability, and system-level integration.

1.2 PRINCIPLES OF DATA CONVERSION

1.2.1 Fundamental processes of analog-to-digital conversion

Being as fundamental as RGB for color images, analog-to-digital conversion also implies the execution of an “SQE” triple fundamental process: Sampling, Quantization, and Encoding (Fig. 1.2). The sampler receives an analog input signal $x(t)$, which is by nature continuous time and can assume any of the infinite values between the negative full scale range $-FSR/2$, and the positive full scale range $+FSR/2$, and transforms it into a periodic sequence of samples $x_n = x(nT)$ with period T , with n an integer. T is the sampling period and $f_s = 1/T$ the sampling frequency.

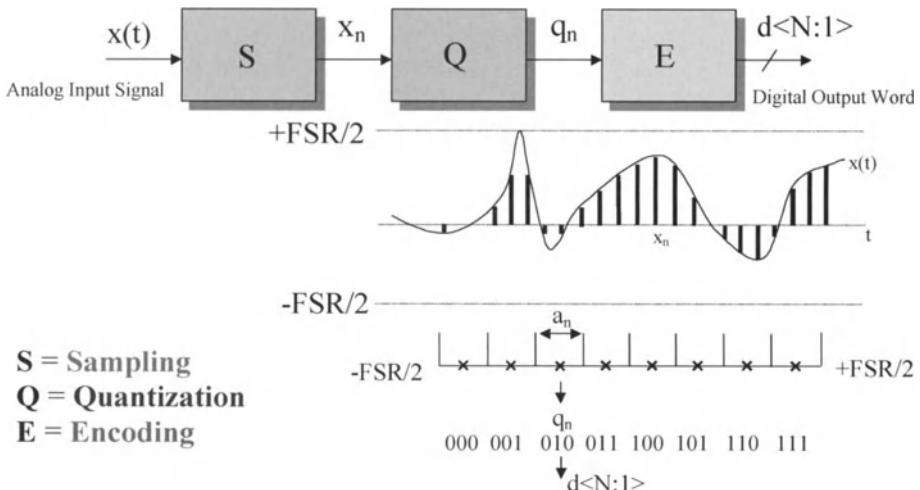


Figure 1.2: The fundamental processes of analog-to-digital conversion:
sampling, quantization, encoding

The sampling process performs a quantization of the signal in the time domain, without affecting the amplitudes of the signal. At this stage, analog samples are available within the data converter, but it is not possible yet to digitally encode these samples, as an infinite number of bits would be required to represent each one. For this reason, a quantizer is needed. The quantizer divides the full scale range $[-FSR/2, +FSR/2]$ into $n=2^N$ intervals, each of amplitude a_n , and associates to the infinite number of analog values within each interval, only a single value q_n . Due to the quantization of the amplitudes of the input analog signal, it is finally possible to associate a digital code with finite length to each sample q_n . This happens in the encoder, that generates the output digital word $d<N:1>$ for each received quantized sample q_n . For a quantizer with $n=2^N$ levels,

N bits are necessary and sufficient to create the output word $d<N:1>$, and the analog-to-digital converter is said to be an “N-bit converter”.

1.2.2 Quantization

It should be noted that quantization is a non-reversible process degrading the quality of the input signal, this being mathematically and physically reflected by the introduction of a quantization noise. In Fig. 1.2 it has been assumed that the quantizer follows a uniform quantization law, which means that all intervals have the same amplitude. A uniform quantization law is generally used in data converters but, considering that the goal of the quantizer and its quantization law should be to limit the level of degradation introduced, it may be not the best choice in some cases. If the characteristics and the statistics of the signal to be converted are well known, as it happens in some applications (voice, video, digital modulated signals), then it can be possible to apply a specific quantization law, which tends to minimize the quantization noise power.

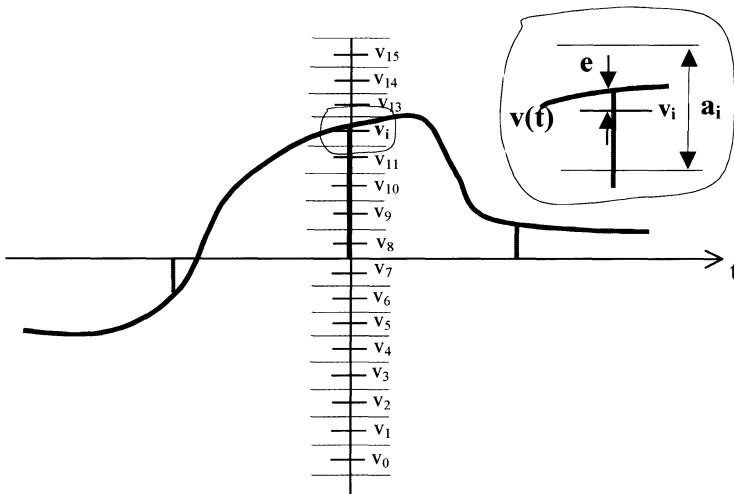


Figure 1.3: Quantization noise general signal analysis

Let us carry out a more general analysis on quantization noise [4] with the aid of Fig. 1.3. For an input signal $v(t)$, the figure shows that after sampling and quantization, the “transmitted” sample is not anymore $v(t)$, but the quantized value v_i , equal to superposing on $v(t)$ a noise $-e$ (quantization noise):

$$\text{Quantization Noise} = e(t) = v(t) - v_i \quad (1.1)$$

To perform an analysis which is as much general as possible, we have to assume that the signal $v(t)$ can have different probabilities to fall into different quantization intervals, but we can also assume that, if all intervals are small

enough, all the values within the i -th interval have the same probability of occurrence:

$$P(v \text{ in } i) \approx f_v(v_i) * a_i \quad (1.2)$$

(assuming that the interval is small, the integral is avoided). According to what just said, we will have in general:

$$P(v \text{ in } i) \approx f_v(v_i) * a_i \quad ? \quad P(v \text{ in } j) \approx f_v(v_j) * a_j \quad (1.3)$$

The two probabilities can be different because the probability density function values f_v are different at v_i and v_j , or because the two interval amplitudes a_i and a_j are different, or the two.

On the other hand, once $v(t)$ is in the i -th interval, we have assumed that $v(t)$ can assume all interval values around v_i with the same probability, then the same applies to $e(t) = v(t) - v_i$:

$$f_e(e/v_{in}i) = 1/a_i \quad (1.4)$$

We can now calculate the variance s_i^2 of the statistical variable “quantization noise e ” of the signal v , conditioned to the event “ v is in i -th interval”:

$$\sigma_i^2 = \int_{v_i - ai/2}^{v_i + ai/2} (v - v_i)^2 f_v(v/v_{in}i) dv \quad (1.5)$$

With the change of variable $e = v - v_i$:

$$\sigma_i^2 = \int_{-ai/2}^{+ai/2} (e - v_i)^2 f_{v-vi}(e/v_{in}i) d(e - v_i) \quad (1.6)$$

$$\sigma_i^2 = \int_{-ai/2}^{ai/2} e^2 f_e(e/v_{in}i) de = \int_{-ai/2}^{ai/2} e^2 / a_i de = \frac{a_i^2}{12} \quad (1.7)$$

Assuming that quantization noise is an ergodic process, which means that statistical averages equal temporal averages, the quantization noise power (equal to the variance s_i^2) is $a_i^2/12$, and the quantization noise rms value (equal to the standard deviation s_i) is $a_i/\sqrt{12}$:

$$s_i^2 = \text{quantization noise power of } e \text{ when } v \text{ is in interval } i = a_i^2 / 12 \quad (1.8)$$

$$s_i = \text{quantization rms value of } e \text{ when } v \text{ is in interval } i = a_i / \sqrt{12} \quad (1.9)$$

Finally, we can calculate the total quantization noise power by adding all the noise contributions:

$$\sigma^2 = \sum_1^{2^N} P(v_{in}i) \frac{a_i^2}{12} = \sum_1^{2^N} f_v(v_i) a_i \frac{a_i^2}{12} \quad (1.10)$$

Eq. (1.10) and also Eq. (1.3) show that it is possible to minimize the quantization noise power by creating a higher density of quantization levels where the signal has a higher probability to occur (nonuniform quantization). However, for a signal $v(t)$ having a constant distribution of amplitudes, Eq. (1.10) shows that uniform quantization, with $a_i = a = V_{LSB}$ for all 2^N intervals, is the optimal choice giving:

$$\sigma^2 = \sum_1^{2^N} P(v_{in}i) \frac{a^2}{12} = \frac{a^2}{12} \sum_1^{2^N} P(v_{in}i) = \frac{V_{LSB}^2}{12} \sum_1^{2^N} \frac{1}{2^N} = \frac{V_{LSB}^2}{12} \quad (1.11)$$

From Eq. (1.11) it turns out that: “Quantization noise is inversely proportional to the square number of quantization levels: to reduce it, converter resolution must increase”. Coming directly from Eq. (1.11) is the expression of the peak SNR for a sinewave input signal:

$$SNR = 20 \log \frac{FSR / (2\sqrt{2})}{V_{LSB} / \sqrt{12}} = 20 \log(2^N \sqrt{3/2}) = 6.02N + 1.76dB \quad (1.12)$$

Very often Eq. (1.12) is considered as a general result; on the contrary, the initial assumptions to get it should be not neglected: peak SNR for an A/D converter with uniform quantization, and expression applicable only for an input sinewave.

It is worthwhile to conclude the analysis on the quantization process, by considering some practical examples of non-uniform quantization. The previous equation shows the expression of SNR for a perfect sinewave, whose peaks reach the full scale of a uniform quantizer. Even if mathematical expressions can help to get an estimate of the performance, the question now is when it can happen that an ADC has to convert such a kind of ideal signal at its input. Dealing with “real word” signals can be very different from dealing with sinewaves. Considering the real case of a voice signal, its statistical properties have been studied in detail. Voice signals have a quite large crest factor (ratio between peak power and mean power), in the order of 20dB, but the peaks of power occur only for 0.01% of the total time. This means that the distribution of amplitudes, mathematically the probability density function, is not constant. Several empirical probability density functions have been given for the voice signal: exponential, gaussian, sinusoidal,...

According to this information, and the analysis performed before, it is easy to be convinced that the uniform quantization law is not practical for voice signals:

- i) given the non-constant distribution of amplitudes, uniform quantization is not the law which minimizes the quantization noise power. In this case, interval amplitudes must be decreased where the signal probability is higher, and must be decreased where lower.

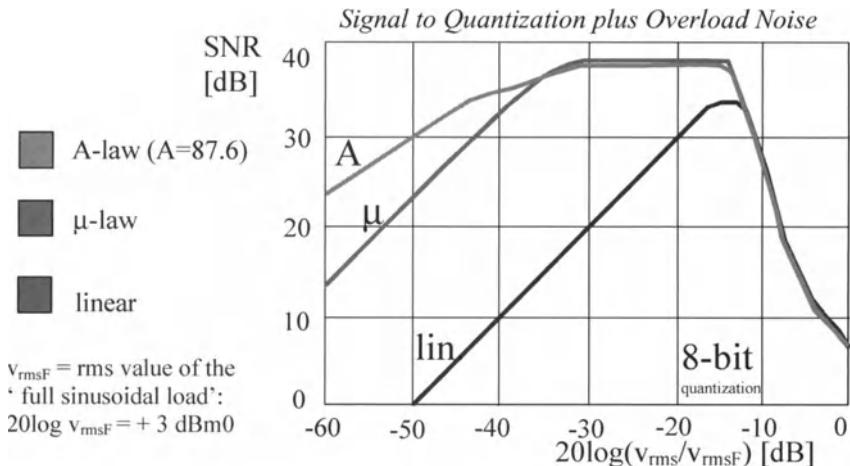


Figure 1.4: Nonuniform quantization applied to voice signals

ii) Eq. (1.11) showed that for a uniform quantization law, and a fixed full scale, arising from the maximum signal amplitude, the quantization noise power is constant. This means that the peak SNR is obtained when the signal reaches the full scale, but when the signal amplitudes decreases, the SNR decreases linearly with it. Unfortunately, voice signals in real applications can have very different and unpredictable power levels, causing a problem of highly variable SNR (Fig. 1.4). From Fig. 1.4 we see that the non-uniform quantizations μ -Law and A-Law, not only increases SNR with respect to uniform quantization, but also allow, in quite large interval of amplitudes, a more independent SNR as a function of the input level. The sharp decrease of SNR from -10dB to 0dB happens because the quantizer can have very few quantization levels near the full scale, or at the limit because the full scale can be made smaller than the peak amplitude. These are real life tradeoffs. Choosing the full scale smaller than the peak signal, which has a very low probability to occur, may allow, for a given resolution, to decrease the LSB then the noise power. On the other hand, when the signal overflows the full scale range (overload condition), the quantizer clamps it, with an effect equivalent to a noise, called "overload noise". It is worth observing that the G.711 Recommendation defines the level of the "full sinusoidal load", namely a sinusoid whose amplitude corresponds exactly to the quantizer full scale, as +3dBm0 with respect to a reference level of 0dBm0, where all the electrical characteristics of the converter are normally specified. This definition suggests that a quantizer has never to be operated, in typical signal conditions, near its full scale and that, also in these conditions, a peak signal may have a probability of overloading the data converter.

1.2.3 Sampling, downsampling and oversampling

The Shannon Theorem is at the basis of sampling [5], and its conclusion is the most fundamental principle exploited by the “full Nyquist-rate” converters. It says: ‘If a signal $x(t)$ has a *limited bandwidth* (-BW, BW), it can be univocally determined by its samples $x(nT)$ if the *Sampling Frequency is at least twice the Bandwidth*: $f_s = 1/T \geq f_n = 2 \cdot \text{BW}$. f_n is called Nyquist frequency.

We should note that:

- 1) Limited bandwidth is a necessary, but not sufficient condition;
- 2) Sampling frequency at least twice the bandwidth is a sufficient, but not necessary condition.

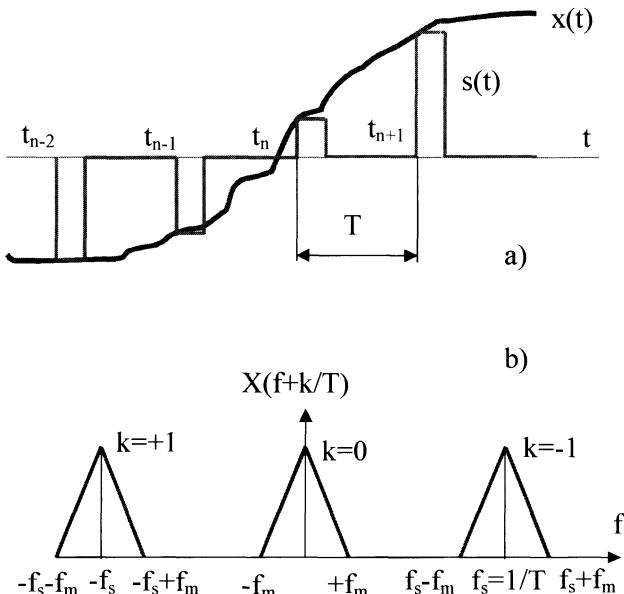


Figure 1.5: Real sampling of a signal $x(t)$ with a ‘carrier’ pulse $g(t)$

In Fig. 1.5 a), an input signal $x(t)$ is sampled and held by a pulse $g(t)$. The resulting signal is $s(t)$:

$$s(t) = \sum_n x(nT)g(t-nT) \quad (1.13)$$

This waveform resembles what is delivered by real sampling circuits like sample-and-holds: $s(t)$ is a succession of pulses $g(t-nT)$, where $g(t)$ is the ‘carrier’ pulse, modulated in amplitude by the initial signal $x(t)$ containing the information. This kind of signal has the characteristics of the so-called PAM (Pulse Amplitude Modulated) signal.

The Fourier Transform of $s(t)$ gives the following expression (we remind that the Fourier Transform by definition has both positive and negative frequencies):

$$S(f) = \frac{1}{T} G(f) \sum_k X(f + \frac{k}{T}) \quad (1.14)$$

Fig. 1.5b) shows the spectral content of the succession of $X(f+k/T)$. Given the fact that $X(f)$ is band-limited, also $S(f)$ will be band-limited, with a periodicity of $f_s=1/T$. By quick inspection of Fig. 1.5b), it is possible to see that if the $X(f+k/T)$ succession of pulses does not overlap in frequency, it is possible to filter and recover the “baseband” pulse $X(f)$, then $x(t)$. The non-overlap condition is:

$$f_m < f_s - f_m ? \quad f_s > 2f_m \quad (1.15)$$

which is the sufficient condition of the Shannon Theorem to recover the original signal. How to recover the signal is quite simple, as Fig. 1.6 shows.

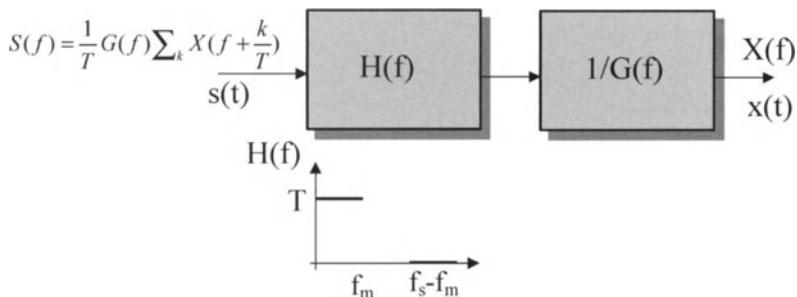


Figure 1.6: Reconstruction of $x(t)$

The signal is firstly processed by a low-pass filter $H(f)$, which is flat from dc to f_m , and has infinite attenuation from f_s-f_m , and after equalized with a transfer function $1/G(f)$, which removes the effects of real sampling. It is easy to argue that the reconstruction filter is simplified, if the Shannon condition (1.15) is verified with enough margin.

Another way of sampling an analog signal, which is also related to analog-to-digital converters and their applications in telecom, is downsampling. It is a special sampling technique that still permits to fully recover the original continuous time signal, but that does not satisfy the Shannon Theorem. Assume that the signal $x(t)$ is not only band-limited, necessary condition as per Shannon theorem, but also band-pass. Under this assumption, it is still possible to recover the original signal $x(t)$, even if we downsample it with respect to Nyquist frequency, $f_s < 2f_m$. Fig. 1.7 shows the effect of sampling $x(t)$ with a sampling frequency $f_s < f_i < 2f_m$. Given the periodical nature of the sampled signal, with frequency shifts multiple of f_s , it is still possible to filter and recover a frequency translated replica of $X(f)$. This effect of downsampling, which produces a shift in frequency of the signal, is often exploited in communications and especially wireless applications, to downconvert and RF or IF signal to a lower frequency.

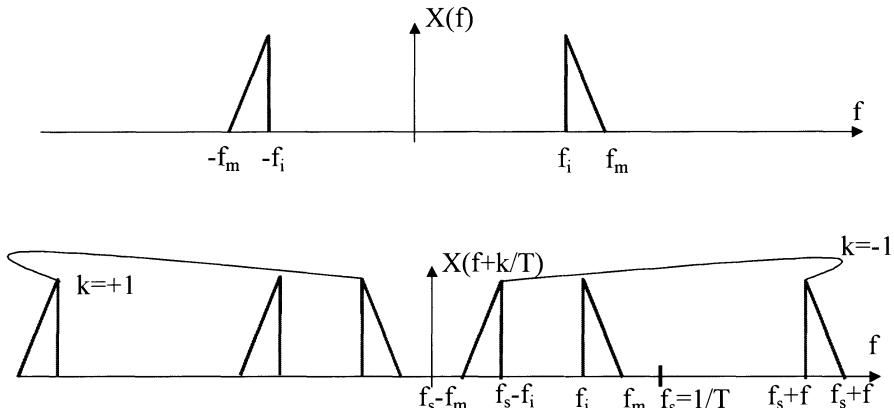


Figure 1.7: Downsampling

The operation is similar to the one performed by a mixer, but care should be taken, since the S&H and the mixer do not operate physically in the same manner. As Fig. 1.8 depicts, one signal is at an IF frequency of 52 MHz, and it is downsampled at $f_s = 13$ MHz with the intention to down-convert it to baseband. Signal replicas are obtained at 39 MHz, 26 MHz, 13 MHz, and baseband. If before downsampling the IF signal is not pre-filtered, the noise floor will be folded four times in the baseband, causing a worsening of the SNR. The example demonstrates that in downsampling application, it is mandatory to filter before, to not degrade the signal.

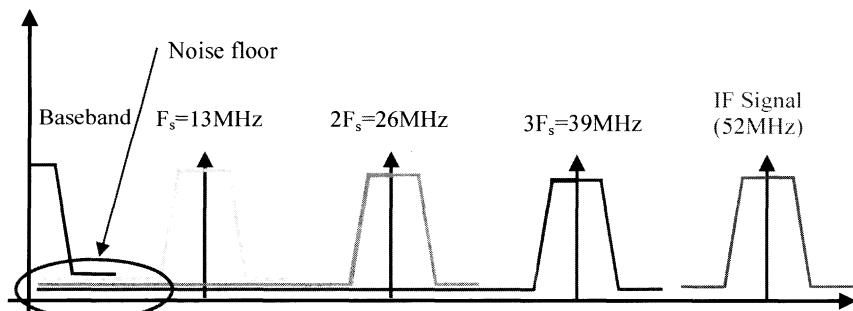


Figure 1.8: Effect of downsampling without pre-filtering the IF signal

Fig. 1.9 shows one application, where an ADC is used in a receive demodulating channel, with its S&H operating in downsampling mode [6]. The RF signal is first down-converted through a mixer to a first IF frequency of 81.25 MHz. After a filtering of the IF signal, to avoid aliasing of the noise floor, the S&H of the A/D converter downsamples the incoming signal at a sampling frequency of $f_s = 13$ MHz. The signal, after downsampling, looks like the one shown in Fig. 1.8, with the lowest IF component present at $81.25 \text{ MHz} - 6f_s = 3.25 \text{ MHz}$.

The signal is digitized by the A/D converter, and further downconverted to baseband with in-phase (I) and quadrature (Q) components by means of a digital oscillator at 3.25 MHz and quadrature demodulator.

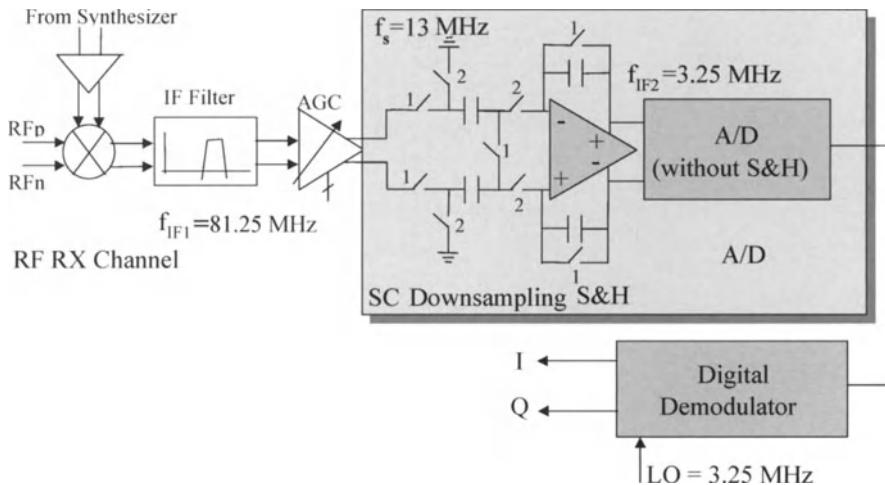


Figure 1.9: An RX demodulating channel using an A/D converter in downsampling mode

Oversampling is another technique very commonly used in data converters. Even if it is largely described in the chapters of this book dedicated to oversampled converters, we will briefly mention it here, since Nyquist-rate converters are in reality always slightly oversampled.

Fig. 1.10 shows a complete analog-to-digital conversion channel, with an analog antialiasing filter AAF, whose transfer function is $H_a(f)$, the ADC, and a digital filter, whose transfer function is $H_d(f)$. According to the Shannon theorem, sampling at Nyquist rate $f_s=2f_m$ fully preserves in principle the characteristics of the signal. Anyway, under this extreme condition, due to the fact that signal spectrum replicas will practically “touch” each other and, as we introduced for downsampling, the original signal is not rigorously band-limited (noise and spurious tones are present out of the band of interest), a very sharp, practically unfeasible analog antialiasing filter, with zero transition band between the passband and the stopband, will be necessary to preserve the maximum achievable signal-to-quantization noise ratio. The example of Fig. 1.10, where an oversampling of 4 is performed, $f_s = 4(2f_m) = \text{OSR} \cdot f_N$, appears less critical for the antialiasing filter, which has now a non zero transition band. For a better understanding, the figure shows for the positive frequencies (but the same must apply for the negative ones) the spectrum replicas of both the signal and the AAF response after sampling. It is evident how no additional spurious noise falls in the band of interest, due to the infinite attenuation provided by the AAF. Antialiasing filter design feasibility is not the only advantage arising from oversampling.

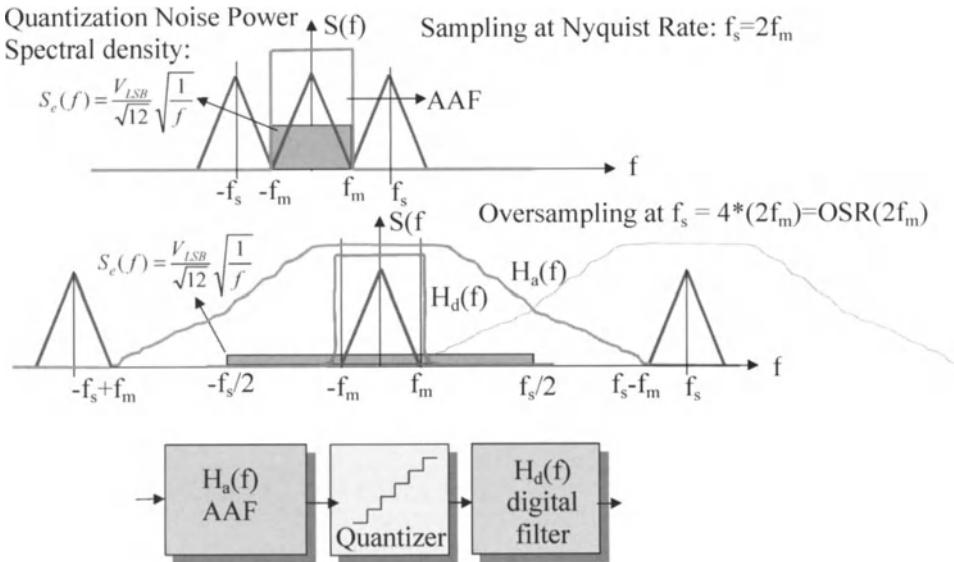


Fig. 1.10: Oversampling benefits versus Nyquist sampling

Assume that, under certain conditions for the input signal, which are normally satisfied, quantization noise has a constant power spectral density between $-f_s/2$ and $+f_s/2$, and it is zero elsewhere, as depicted in Fig. 1.10. Moreover, according to Eq. (1.11), the quantization noise power spectral density assumes the following expression, where different from zero:

$$S_e(f) = \frac{V_{LSB}}{\sqrt{12}} \sqrt{\frac{1}{f}} \quad (1.16)$$

Oversampling the signal will allow the quantization noise to be spread over a larger bandwidth, and in particular out of the bandwidth of the signal. Once in the digital domain, the digital filter $H_d(f)$ will remove the quantization noise out of the maximum signal frequency, signal-to-noise ratio will consequently be increased.

Eq. (1.11) and (1.16) show that, for each doubling of the sampling frequency, the SNR of the quantizer is improved by 3dB, corresponding to an increased resolution of half-bit.

In conclusion, a slight oversampling is normally applied also to Nyquist rate converters, because it allows antialiasing filtering requirements to be relaxed, and also improves the SNR.

1.3 ARCHITECTURES OF NYQUIST RATE CONVERTERS

1.3.1 ADC classification

A convenient way to classify all data converters, is to group them into different categories, which differ in terms of conversion speed. Then, a way to rapidly inspect the speed of each converter is to see how many clock cycles are used to perform a single conversion. The numbers of clock cycles per conversion should be not confused with the number of samples per conversion.

Three main categories are identified:

- 1) Converters using an exponential number of cycles, in the order of 2^N , where N is the converter resolution. The integrating dual ramp and incremental ADC, which can offer very high resolution (16-bit or more) are part of this category.

Type	Clock Cycles/ Conversion	Family
<u>VERY FAST SPEED - MEDIUM, LOW RESOLUTION</u>		
Folding	1	(full) Nyquist
Full Flash	1	(full) Nyquist
Pipeline	1-2	(full) Nyquist
Two-step Flash	2	(full) Nyquist
<u>MEDIUM, FAST SPEED - HIGH, MEDIUM RESOLUTION</u>		
Successive. Approximation	$\sim N$	Nyquist
Algorithmic	$\sim m * N$	Nyquist
Sigma-Delta	$m * N < 2^k < 2^N$	Oversampled
<u>SLOW SPEED – VERY HIGH RESOLUTION</u>		
Incremental	$[2^N, 2^{N+1}]$	Nyquist
Integrating dual ramp	2^{N+1}	Nyquist

Fig. 1.11: Data converters classification versus number of clock cycles per conversion. Only the sigma-delta has to be considered out of the list of Nyquist-rate converters.

- 2) A very wide category of converters of medium-high speed and high-medium resolution, which use a number of clock cycles still exponential 2^k , with k somewhat lower than N, like sigma-delta converters, others which use $m * N$ clock cycles, arising from m clock cycles used to resolve each bit, like the algorithmic converters, and finally the successive approximation converters, which use normally around N clock cycles, with one clock cycle per bit.
- 3) The last category of highest speed converters, which use just 1-2 clock cycles to perform a conversion. Here we find the two-step flash, the full flash, the pipeline, and the folding ADC converters.

The terms “high”, “medium”, “low” resolution are purely indicative, and must be interpreted in a flexible way. For example, a pipeline or two-step flash “medium” resolution ADC can be in the order of up to 10-bits, anyway if self-calibration techniques are deployed, its resolution may increase to 12-bits or more. In this paragraph we will make an overview of the fundamental architectures, but it is not uncommon to find data converters exploiting a combination of the ones listed in Fig. 1.11.

The classification of ADCs between the two big families of Oversampled or Nyquist rate ones, is not obvious. We can categorize as Nyquist-rate converter, an ADC capable to operate under Nyquist condition, namely with a sampling close to half of the maximum input frequency.

All the converters listed in Fig. 1.11, apart from the sigma-delta, are suitable to operate in this way. The sigma-delta ADC, due to its structure, is quickly losing its performance, if some level of oversampling is not applied. For some sigma-delta architectures, it is not mandatory to keep high oversampling ratios to achieve high performance. In fact, sigma-delta converters, even if commonly defined as oversampled converters, exploit the benefits of combining oversampling with quantization noise shaping.

On the other hand, as introduced in previous chapter, Nyquist rate converters, whenever possible, are slightly oversampled. Only the fastest Nyquist-rate ADCs in the category of 1-2 clock cycles (flash, pipeline, folding), are typically used in extreme sampling condition, to not lose any speed performance. For this reason, they are also defined “full Nyquist-rate” converters.

1.3.2 The integrating dual ramp ADC

The integrating dual ramp ADC is a very accurate but slow one. If N is the resolution of the ADC, conversion time can reach $2^{N+1}/f_{\text{clk}}$. For example, assuming $f_{\text{clk}} = 1\text{MHz}$ and $N=12\text{-bit}$, conversion time will take around 4ms.

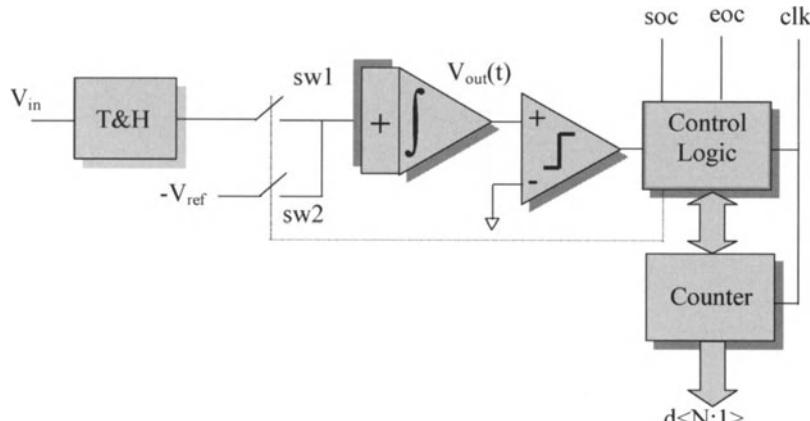


Fig. 1.12: Block Diagram of the integrating dual ramp ADC

Higher resolutions, up to 16-bit or more, with longer conversion times, are possible. Fig. 1.12 shows the block diagram of the integrating dual ramp ADC.

When $\text{soc}=1$ is asserted, conversion starts, switch $\text{sw}1$ is closed and the input signal V_{in} is integrated, V_{out} ramps up and the counter, which was reset at beginning of conversion, starts to count. This first phase, integration of V_{in} , stops when the counter counts 2^N , then it overflows and generates a control signal which opens switch $\text{sw}1$ and closes switch $\text{sw}2$. The first integration phase has a fixed time duration T_1 , and the integrator output voltage $V_{\text{out}}(T_1)$ is proportional to the input voltage, this meaning the ramp slope is proportional to V_{in} , Fig. 1.13. In the second integration phase, the counter is reset again and V_{out} ramps down at a constant slope, due to the fact that the integrator input is at a constant voltage V_{ref} . The second integration phase has a variable time duration T_2 , proportional to V_{in} via $V_{\text{out}}(T_1)$, and stops when the integrator output reaches the comparator threshold voltage: at this time the eoc (end of conversion) is flagged to one and the counter is read out. It is easily demonstrated that the counter final state is the digital representation of V_{in} . After the first integration at constant T_1 , the output voltage is:

$$V_{\text{out}}(T_1) = \frac{1}{RC} \int_0^{T_1} V_{\text{in}} d\theta = \frac{V_{\text{in}} T_1}{RC} \quad (1.17)$$

where RC is the time constant of the integrator.

The conversion is stopped at the end of phase 2 when V_{out} reaches the ground voltage:

$$V_{\text{out}}(T_1 + T_2) = V(T_1) - \frac{1}{RC} \int_{T_1}^{T_1+T_2} V_{\text{ref}} d\theta = \frac{V_{\text{in}} T_1}{RC} - \frac{V_{\text{ref}} T_2}{RC} = 0 \Rightarrow \quad (1.18)$$

$$\Rightarrow T_2 = T_1 \frac{V_{\text{in}}}{V_{\text{ref}}} \quad (1.19)$$

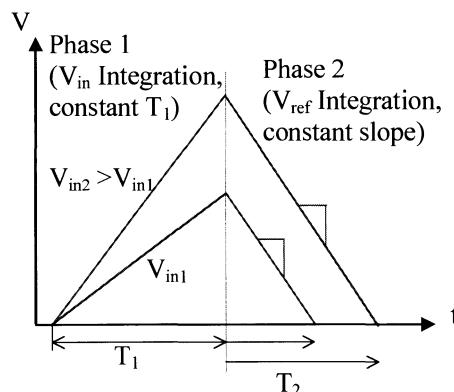


Fig. 1.13. Principle of operation of the integrating dual ramp ADC.

Considering that $T_1=2^N/f_{\text{clk}}$, $T_2=d<N:1>/f_{\text{clk}}$ with Eq. (1.19), we obtain:

$$d<N:1> = 2^N \cdot V_{\text{in}} / V_{\text{ref}} \quad (1.20)$$

It is worth noting that:

- 1) $d<N:1>$ output word is not dependent on f_{clk} accuracy, since the architecture is based on a ratio of counts, more than an absolute number of counts. This is peculiar of the dual ramp integrating ADC, instead of the single ramp integrating ADC, which performs just an absolute number of counts and does show sensitivity to master clock accuracy.
- 2) The converter accuracy depends directly on the V_{ref} accuracy, its stability during the conversion as well as long term stability. Other factors which may affect accuracy are op-amp and comparator offset voltages, noise and other non-idealities.
- 3) Any periodical noise, like 50Hz power line noise, will be rejected if T_1 is an integer multiple of the noise period. This feature is exploited in instrumentation, data acquisition, physical measurements, where integrating dual ramp converters offer the best performance.

1.3.3 The incremental ADC

The incremental ADC is a less well-known, but yet powerful type of slow speed, very-high resolution converter [7,8]. It is interesting to study the principle of operation by correlating it with its two nearest companions which are, according to Fig. 1.11, the integrating dual ramp and the sigma-delta. Actually, the incremental ADC has i) the principle of operation of an integrating converter, as its name suggests, and ii) an hardware implementation very similar to a first order delta-sigma modulator (Fig. 1.14), differing from it for very few details.

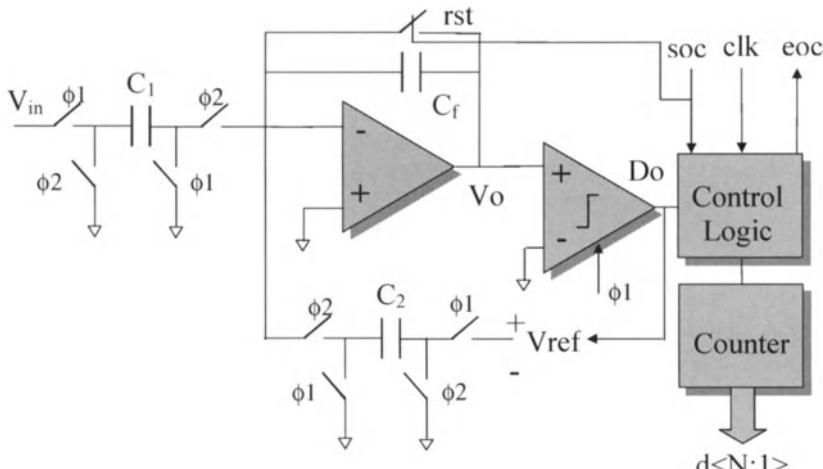


Fig. 1.14: Block diagram of the incremental ADC with circuit details.

The incremental ADC takes $n=2^N$ cycles to perform a conversion, so it is faster than the dual ramp, for the same resolution, by a factor of two. The hardware is much reduced: one integrator, for example of switched capacitor type, one comparator, one counter and digital control logic. At beginning of conversion, stated by soc asserted to one, the integrator is reset. The residue output of the integrator at time instant t_1 , after a first integration, is:

$$Vo(1) = \frac{C_1}{C_f} V_{in}(0) - \frac{C_2}{C_f} V_{ref} Do(0) + Vo(0) \quad (1.21)$$

Similarly, the residue output at instant t_2 , second integration cycle, becomes:

$$\begin{aligned} Vo(2) &= \frac{C_1}{C_f} V_{in}(1) - \frac{C_2}{C_f} V_{ref} Do(1) + Vo(1) = \\ &= \frac{C_1}{C_f} (V_{in}(1) + V_{in}(0)) - \frac{C_2}{C_f} V_{ref} (Do(1) + Do(0)) + Vo(0) \end{aligned} \quad (1.22)$$

And finally, the residue output at the last integration cycle t_n , assumes the following expression:

$$Vo(n) = \frac{C_1}{C_f} \sum_0^{n-1} V_{in}(i) - \frac{C_2}{C_f} V_{ref} \sum_0^{n-1} Do(i) + Vo(0) \quad (1.23)$$

Let's re-arrange Eq. (1.23) by deriving V_{in} , that we assume to be constant during the entire conversion time, and normalizing the expression with $C_1=C_2=C_f$, and $V_{ref}=1$:

$$V_{in} = \frac{1}{n} \sum_0^{n-1} Do(i) + \frac{Vo(n) - Vo(0)}{n} \quad (1.24)$$

Eq. (1.24) fully describes the operation of the incremental ADC. The input voltage has its digital representation in the first term of Eq. (1.24) (normalized to one). The second term is a conversion error. In principle, the worst case errors is obtained when it happens that $Vo(n)=V_{ref}$ and $Vo(0)=-V_{ref}$, giving a worst case error of $2V_{ref}/n$, corresponding to 2 LSB. To reduce it, the residue is nullified at the beginning of conversion, $Vo(0)=0$, by resetting the integrator, thus reducing the maximum worst case error to 1 LSB.

As mentioned at the beginning, the incremental ADC has several commonalties with integrating dual-ramp and sigma-delta ADCs. As will be treated in the chapters dedicated to sigma-delta converters, a sigma-delta modulator is basically composed of an analog modulator loop and a digital low-pass filter. These two blocks can be recognized also in the incremental ADC. While in sigma-delta, specific digital decimation filters are employed, performing an averaging of the input signal, in the case of incremental ADC, where the signal is assumed to be static, the counter still plays the role of a rudimentary rectangular "brickwall" filter.

In the normal use of a sigma-delta ADC with digital decimation filters like comb filters, it is implicitly assumed that the quantizer is "busy", this meaning that the input signal spans, with almost equal probabilities, all the possible amplitudes. Given this assumption, the sigma-delta converter can attain its

theoretical maximum dynamic performance in terms of SNR. On the contrary, if the same sigma-delta is used to convert static dc signals, it is possible to demonstrate that SNR will degrade by some quantity, due to an increase of the noise floor and creation of idle tones. In these conditions, it could be demonstrated also that the accuracy of the sigma-delta would be the same of an incremental ADC, this meaning that a simple rectangular filter (counter) can replace a more complex digital filter [9].

We conclude that the incremental ADC is an optimized architecture for dc signals, whilst the sigma-delta is a full functional, but not optimized architecture for such signals. The difference of operation between the two should be definitively clarified: the sigma-delta is a free-running converter based on oversampling and quantization noise shaping, making use of digital decimation filtering, whilst the incremental ADC is an integrating converter type, whose integrator is reset at each beginning of conversion. As we showed in Fig 1.11, despite the fact that sigma-delta can have high oversampling factors, the number of clock cycles per sample is typically lower than the incremental ADC.

1.3.4 Sigma-delta ADC versus Nyquist rate ADCs

The sigma-delta is a converter type that fully exploits the benefits of both oversampling and quantization noise shaping. Combining these two effects in different manner, leads to a large number of sigma-delta architectures. Very few basic concepts will be mentioned here, just to carry out a comparison with the Nyquist rate converters.

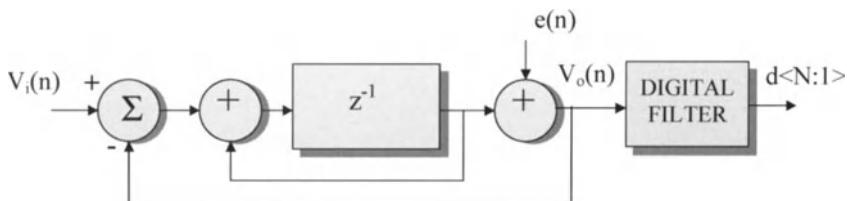


Fig. 1.15: A first order ($L=1$), 1-bit delta-sigma A/D converter

Fig. 1.15 shows the simplest form of sigma-delta ADC, a first-order, noise shaped 1-bit ADC. By simple calculation we obtain for the transfer function:

$$V_o(n) = z^{-1}V_{in}(n) + (1-z^{-1})e(n) \quad (1.25)$$

The achievable peak SNR is:

$$\text{SNR}_{\max} = [(+6.02N+1.76) - 5.17 + 30\log\text{OSR}] \text{ dB} \quad (1.26)$$

Each doubling of f_s implies $3 \text{ dB} + L*6 \text{ dB} = 9 \text{ dB}$ SNR increase (3 dB due to OSR, 6 dB due to noise shaping).

The advantages of a joint use of oversampling, noise shaping, and $N=1$ -bit analog-to-digital conversion appear evident. Assume this converter clocked at

1.024 MHz for converting a voice signal, whose Nyquist frequency is 8 kHz. Given the high oversampling ratio (OSR=128), the achievable SNR is around 65dB. Thanks to the combined use of oversampling and noise shaping, around 10-bit effective performance is achieved, in conjunction with an almost ideal linearity, coming from the utilization of an intrinsically linear 1-bit quantizer. A second order modulator, as shown in Fig. 1.16, further enhances the benefits of noise shaping.

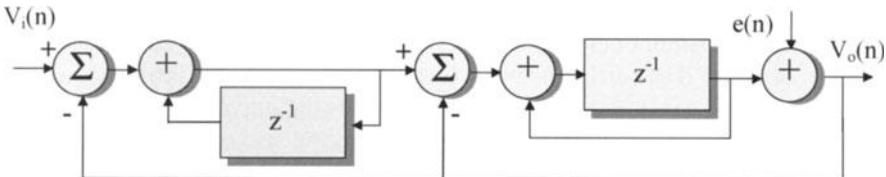


Fig. 1.16: A second order ($L=2$), 1-bit delta-sigma A/D converter

For the second-order 1-bit modulator the transfer function becomes:

$$V_o(n) = z^{-1}V_{in}(n) + (1-z^{-1})^2e(n) \quad (1.27)$$

and the peak SNR:

$$\text{SNR}_{\max} = [(+6.02N+1.76) - 12.9 + 50\log\text{OSR}] \text{ dB} \quad (1.28)$$

In case of second order 1-bit sigma-delta, each doubling of f_s implies $3 \text{ dB} + L * 6 \text{ dB} = 15 \text{ dB}$ SNR increase (3 dB due to OSR, 12 dB due to noise shaping). Coming back to the example of the voice signal, we can now obtain $\text{SNR}_{\max} \sim 99 \text{ dB}$. This performance is impressive, and it explains by itself why the second order sigma-delta is so widely used in voice applications, and similar ones. The ideally lowest possible distortion, arising from the use of a 1-bit quantizer (comparator) should also be considered: modulator loops with 1-bit quantization do not require tight matching of analog components (resistors, capacitors, or transistors), nor accurate characterization and modeling in simulation of these components, so that they can offer, in principle, very good and predictable results.

At this point we may raise the legitimate question: why to use other Nyquist rate converters if the sigma-delta one looks so powerful and convenient?

From an industrial point of view, considerations of feasibility, yield, robustness, and time-to-market, suffice to say that if a 1-bit sigma-delta converter can be identified as a solution to cover a given specification, it should be always retained as the first option.

Unfortunately, the sigma-delta gives its best performance when it can trade-off oversampling with low signal bandwidth. If this trade-off works well with not too fast signals, not the same when the input signal frequency content increases, since it becomes no possible, or simply not convenient, to keep high oversampling ratios (OSR).

Let's consider again the second order modulator performance, in case a 270.83 kHz maximum frequency GSM signal has to be converted. For a 13 MHz master clock, the oversampling ratio is $OSR = 13000/541.66 = 24$, and the theoretical SNR_{max} within a bandwidth of 270.83 kHz is around 64 dB that, compared to (1.12), and considering that Nyquist rate converters will be slightly oversampled, corresponds to a 10-bit Nyquist ADC counterpart. The example shows that as the signal frequency increases, the master clock and the operating frequency of all blocks of the sigma-delta must also increase consequently to guarantee a target SNR. In this condition, the sigma-delta may dissipate a relevant amount of power, and can start to have non-ideal behavior, like limited op-amp slew-rate, clock feedthrough effects, and others. At this point, Nyquist rate converter counterparts not only become available, but can represent a more competitive option.

1.3.5 The successive approximation ADC

The successive approximation ADC is a very popular architecture. It can cover a wide range of resolutions, up to 10-bit or even 12-bit, with careful design, layout and process characterization, and up to 16-bit with self-calibration techniques. Conversion rates can also cover a very wide range, from few kHz up to several MHz. The increase of speed over the “barrier” of 1 Msps has been feasible with the introduction of faster, deep sub-micron technologies: it is possible to find examples where the converter, fabricated in CMOS technology, is clocked at several tens of MHz [10]. On the other hand, it is still a low-power architecture, due to the utilization of a small number of building blocks: a sample-and-hold, a comparator, and a digital-to-analog converter.

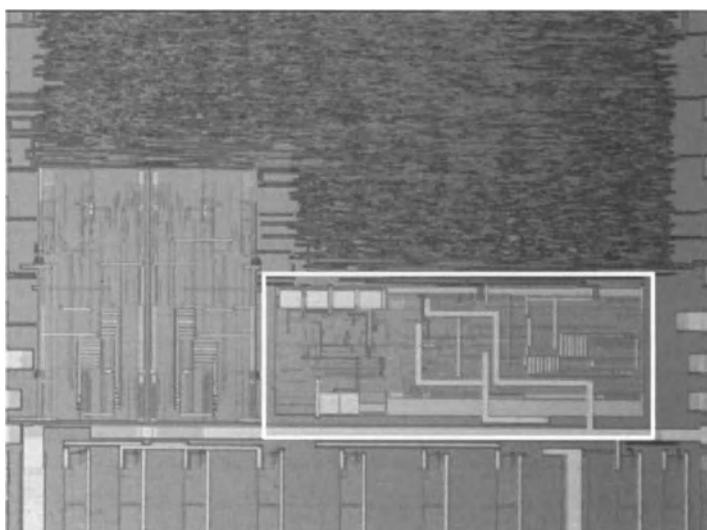


Fig. 1.17: An 8-bit, 100ksps, 100uA, 1.8V supply ADC in Atmel CMOS technology, embedded in a ASIC for remote sensing applications

Due to these features, the successive approximation ADC finds many applications in very different fields: low-speed, micro-power remote sensing, medical applications such as pacemakers, up to higher speeds and resolution like wireless baseband demodulators.

According to Fig. 1.18, the principle of operation is based on the attempt, through “successive approximations”, of reproducing the analog input signal by means of the output of a digital-to-analog converter. This is obtained by minimizing the error signal E at the comparator input, with a binary search algorithm. Once the binary search algorithm is completed, the error between the input and the DAC results minimized, implying that the digital word $b<N:1>$, driving the DAC input, is the digital representation of the input signal. At this point, the word $b<N:1>$ is transmitted at the output, $d<N:1>=b<N:1>$. The complete conversion takes around N clock cycles, where N is the resolution.

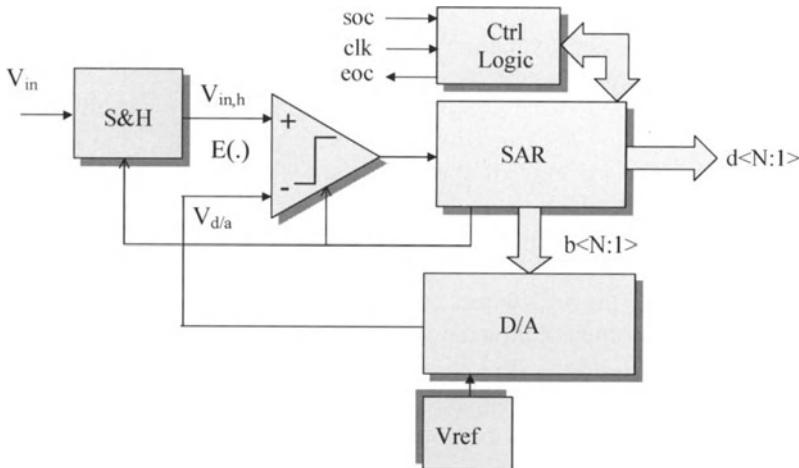


Fig. 1.18: Block diagram of successive approximation ADC.

The binary search algorithm is governed by the successive approximation register (SAR). As already introduced, the unknown input voltage is compared with a “programmable” reference voltage generated by the internal DAC. Voltage steps of $V_{d/a}$ are generated by arbitrarily setting to one, just as an attempt, each bit of the DAC input $b<N:1>$, from the MSB to the LSB. If for each i -th comparison the analog input is higher than DAC output, the bit $b<i>$ is confirmed at one, otherwise is updated to zero.

Even if the converter is properly called successive approximation, it should be considered within a particular class of converters that base their operation on a conversion algorithm. The algorithmic converter, which will be presented in the next paragraph, represents another example of this type of converters. In this sense, the SAR ADC is also an “algorithmic-type” converter.

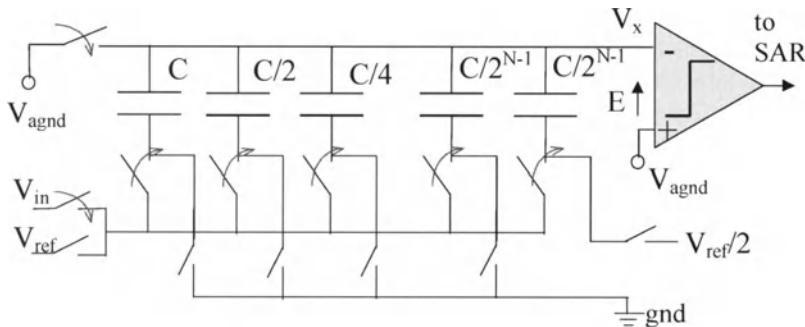


Fig. 1.19: Charge redistribution circuit implementation of a SAR ADC.

It is worth to mention in Fig. 1.19 one of the best-known implementations of the SAR ADC, based on charge redistribution. It is difficult to imagine another converter implementation, which can dissipate less power than this one. The charge redistribution DAC, as described in the following, performs also the sample-and-hold function. Assuming that the design of the reference generator is not critical, which is true for slow speed applications, the consumption of the whole converter is practically the consumption of a comparator. As a consequence, the converter can have a current consumption of few uA or even less than 1uA. The conversion is carried out in three phases:

Phase 1, Sample mode: bottom plates to V_{in} , top plates to $V_{agnd} = V_{ref}/2$:

$$Q = 2C(V_{agnd} - V_{in}) \quad (1.29)$$

Phase 2, hold mode with MSB trial already set: bottom plates to gnd (apart from C to V_{ref} , $b_N=1$)

$$2C(V_{agnd} - V_{in}) = CV_x + C(V_x - V_{ref}) \quad (1.30)$$

$$V_x = V_{ref} - V_{in} \quad (1.31)$$

$$E = V_{ref}/2 - V_{in} \quad (1.32)$$

Phase 3, charge redistribution, bit per bit trials: C_i bottom plate to V_{ref} is $b_i=d_i=1$, to gnd if $b_i=d_i=0$.

Assume for example that V_{in} is at the positive full scale, $V_{in}=V_{ref}$. After sampling and hold with first bit trial to one, the comparator input voltage E is negative and equal to $-V_{ref}/2$. This confirms $b_N=d_N=1$, and the capacitor C bottom plate connected to V_{ref} . After b_{N-1} is asserted to one and the capacitor $C/2$ bottom plate is connected to gnd, which reduces the error E negative voltage to $-V_{ref}/2 + V_{ref}/4$. The conversion continues in the same manner, with all the capacitors bottom plates connected one-by-one to V_{ref} , and the comparator input

error voltage E converging to zero but never being positive. The conversion finishes with $b<N:1>=d<N:1>=111\dots111$.

In conclusion, the SAR ADC is a very versatile converter type, capable of covering a very wide range of specifications in terms of speed, resolution, and consumption. Many circuit implementations are possible, based on resistive, switched-capacitor, current mode, transistor based, M-2M DACs. In any case the accuracy depends on component matching properties, and good process characterization as well as precise simulation models.

1.3.6 The algorithmic ADC

The algorithmic ADC has a dual behavior with respect to the successive approximation ADC. As depicted in Fig. 1.20, the converter operation is still

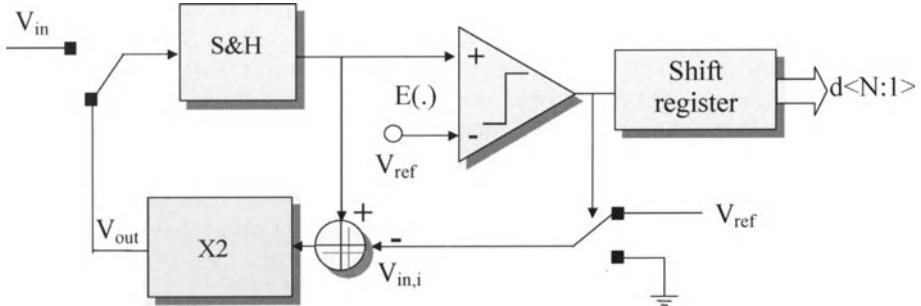


Fig. 1.20: Algorithmic ADC block diagram.

based on a number of comparisons between the input signal and a reference voltage, with the difference that in the algorithmic ADC the reference remains constant, while the input voltage is subject to change. At beginning of conversion, the input signal is directly sample-and-held. After comparison with the reference, if the input signal is higher than V_{ref} , $d<N>$ is set to one, the V_{ref} voltage is subtracted from V_{in} , otherwise $d<N>$ is set to zero and no V_{ref} subtraction is performed. After, the signal is amplified-by-two and redirected at the input of the comparator, where a second cycle starts.

The key function of the converter is the amplify-by-two operation; in some circuit implementations the amplify-by-two and subtraction are both performed by the same hardware. The algorithmic ADC accuracy does not depend directly on component matching, like in SAR ADC or some other Nyquist rate converters, but on the attainable precision on the amplify-by-two and subtract operations.

Several circuit solutions have been found, capable to trade-off precision of the amplify-by-two, with an increased number of clock cycles by a factor m to perform the operation. Looking at the literature, the following results have been achieved for the amplify by two accuracy versus number of clocks m [11]:

- m=1 gain error dependent on component matching, $V_{out,i} = V_{in,i}(1+C_2/C_1)$
- m=2 gain error desensitized but still dependent from component mismatches, $V_{out,i} = V_{in,i}(2 + C^2/2C^2)$, offset compensated
- m=4 gain totally independent on component mismatch, offset compensated
- m=7 gain totally independent on component mismatch, offset compensated, finite gain op-amp independent

The multiply-by-two operation can be performed in an almost ideal manner, if more clock cycles are devoted for converting one bit. In such manner, the algorithmic ADC can reach higher resolutions at expense of slower conversion rates. For instance, at the same clock frequency, the algorithmic ADC results are slower by a factor m with respect to the successive approximation ADC. Typical reported performances are 12-bit at 600ksps [12], up to 14-bit at 10ksps with m=7 [11].

1.3.7 The full flash ADC

The flash ADC, due to the exploitation of a full parallelism, is one of the fastest possible converters, since a conversion is handled within only one clock cycle. Its architecture is attractive because very simple, but it is area consuming and power hungry, and several design trade-offs are necessary as the electrical behavior of each block is investigated in detail (Fig. 1.21).

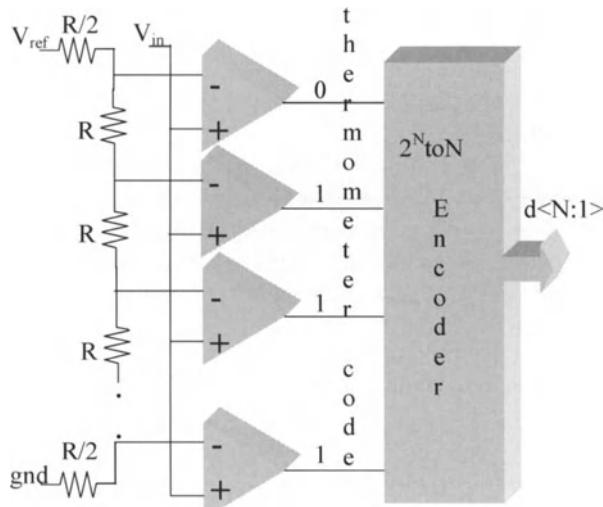


Fig. 1.21: Full flash ADC block diagram.

A resistive ladder, containing 2^N resistors, generates the reference voltages within the full scale range, going to the inverting inputs of the comparators, whilst the input signal is fed into the non-inverting inputs of the 2^{N-1} comparators. At the comparator outputs, a thermometer digital code, proportional to the input signal, is generated, and further converted onto an N-bit code by a 2^N -to-N encoder. An interesting feature of the flash architecture is that an input S&H is not necessary. In fact, the comparators typically use a first amplifier stage, cascaded with a dynamic latch, providing very high gain. The comparators are clocked, and in a first phase the input is sampled and amplified, while in the second the difference between the signal and the reference is instantaneously latched. In practice, assuming that the master clock transition arrives simultaneously on all the latches, the 2^N comparators perform the operation of a distributed sample-and-hold.

There are a number of considerations which limit the maximum resolution for this architecture to roughly N=8-bit. The parallelism implies an exponential increase of area: for 8-bit $2^8=256$ comparators are necessary, whilst for 10-bit they rise up to $2^{10}= 1024$, which is a prohibitive number. The area occupation becomes so significant that makes its deployment not competitive, at least for SOC (System-On-Chip) applications. The increase of number of comparators enhances the input capacitance with the same exponential rule: if from one side the sample-and-hold is not necessary, on the other side a powerful voltage buffer is required to drive the load, and its design becomes impractical if not unfeasible if the capacitive load is excessive. Another limitation can arise from thermal dissipation. Since the flash ADC is used at high speed, the consumption of the comparators is not negligible and power dissipation may be not handled by the IC package over a certain limit. The sizing of the comparator is probably the most critical issue of the design. About this issue, it is worth to note that a random offset of the comparator, has the consequence that the real thermometric code, thus the digital code, is directly affected, producing nonlinearity errors. One possibility could be to reduce the offset by careful design, but this choice implies an increase of the input transistors area, and then of the input capacitance. The other is to perform offset compensation at each cycle, but this often results in loss of conversion speed, caused by the offset compensation, which can be the bottleneck operation in terms of speed; the only way to recover the situation is to increase further the consumption. Other critical design issues, that need to be addressed, exist: loading effect of the resistive ladder, causing nonlinearity, kickback noise, capacitive coupling at the comparator inputs, disturbing the input signal and the reference ladder tap points, clock dispersion, causing non perfect distribute sampling of the input signal. In conclusion, thanks to its simple structure, the flash ADC is an attractive solution in terms of architecture, but resolution should be kept low, for performance and mainly cost considerations.

1.3.8 The two-step flash ADC

The two-step flash ADC architecture (Fig. 1.22) allows the area and power dissipation to be dramatically reduced, at expense of only one conversion each two clock cycles. This implies a sampling frequency reduced by a factor of two, assuming the same speed of the building blocks, and same master clock frequency in the two cases. As the name suggests, in the two-step flash ADC a first coarse conversion of the signal happens, with the MSBs (Most Significant Bits) converted, then the “residue”, namely the difference of the input signal and the analog representation of the coarse MSB conversion (obtained through a DAC), is fine converted, allowing to resolve the LSBs (Less Significant Bits). To understand the advantages in terms of area and dissipation, assume an $N=8$ -bit fast converter implementation with a full flash ADC and a $4+4$ two-step flash. In the first case, $2^N = 256$ comparators have to be used, whilst in the second only $2^{N/2} + 2^{N/2} = 32$ comparators are necessary. Looking more in detail Fig. 1.22, new building blocks appear in the architecture with respect to the full flash: an input S/H, a DAC, the subtractor and residue amplifier. From the master clock frequency, two phase signals ϕ_1 and ϕ_2 , controlling the timings of the different blocks, are generated. Conversion begin with S/H in sample mode; during this phase, which is not active for the other blocks in terms of signal processing, other operations, tending to improve the precision of the converter, can be launched, like comparator offset compensation, residue amplifier gain calibration, and others.

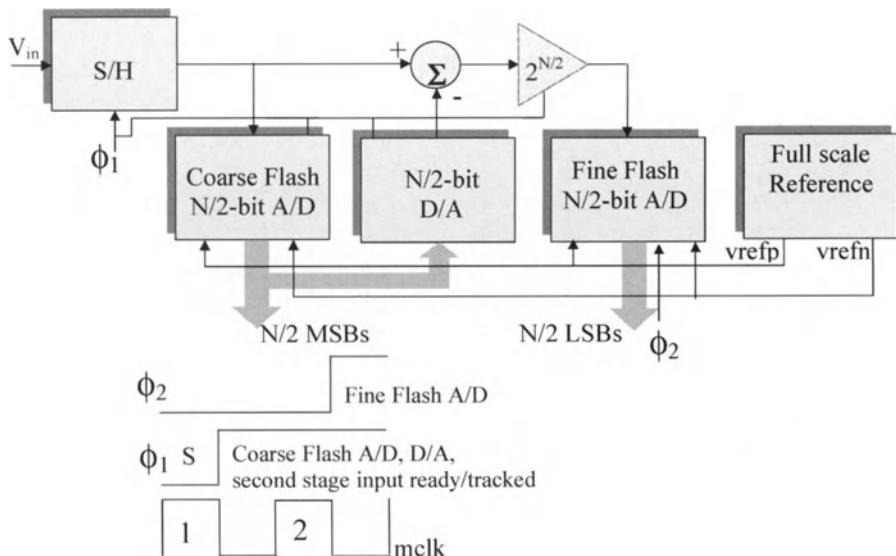


Fig. 1.22: Two-step flash ADC block diagram.

Once ϕ_1 changes its state, the input signal is held and the coarse flash conversion happens. Typically, in one-half master clock period the signal polarity can be amplified in the comparators, in the second the same comparators are latched, and the first $N/2$ MSBs are converted. These $N/2$ MSBs can drive now the input decoder of the DAC, which can settle within one-half period of a master clock. In parallel, since the input signal is available at the output of the S&H, the residue generation and amplification can take place. Finally, the fine flash conversion occurs at the ϕ_2 transition, with a first half master-clock period duration to amplify the signal, and the second to latch the LSBs. In the timing diagram of Fig. 1.22, which is indicative, the conversion begins on the first rising edge of master clock and terminates on the third rising edge with the last operation, (LSBs latched), taking in total two clock periods. The S/H function is necessary for the two-step flash, since the input signal acquisition happens a first time ($\phi_1 = 1$) for the coarse flash conversion, then a second time, with a delay between one-half and one master clock, to generate the residue. If during this double delayed signal acquisition, the signal is not held to offer a stable voltage, big conversion errors can take place, especially if the signal experiences fast excursions during the delay.

As depicted in Fig. 1.22, the residue amplifier allows the use a fine flash quantizer having the same structure and full scale range of the coarse quantizer.

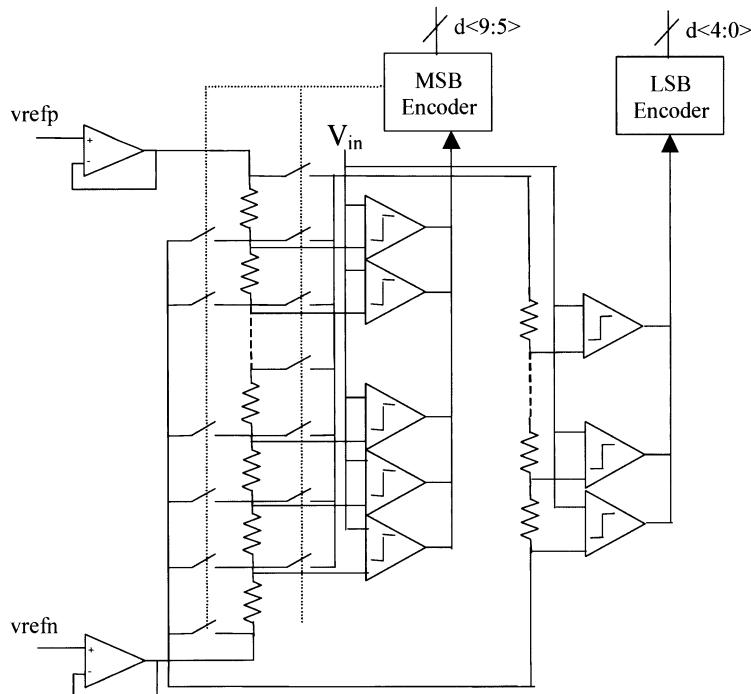


Fig. 1.23: Two-step flash subranging ADC circuit implementation with resistive DAC.

This architectural solution may be convenient or not, depending on a number of circuit details. One hand, the residue amplification allows the fine flash ADC to operate with stronger signals, with consequent relaxed requirements for the design of comparators in terms of offset and speed, on the other hand it assumes a perfect behavior for the residue amplifier, which must be fast enough, and amplify the residue without gain errors.

An alternative solution, that does not require subtractor and residue amplification, is the two-step flash subranging ADC. For a coarse conversion of $N/2$ bits, we can imagine that the full scale range is subdivided into $2^{N/2}$ smaller subranges. Once the coarse flash conversion is performed, the first MSB conversion result indicates in which of the $2^{N/2}$ subranges the input signal lies. At this instant, the full scale of the fine flash quantizer is programmed to the identified subrange, and the conversion of the last $N/2$ bits can take place.

Fig. 1.23 illustrates a circuit implementation, based on a two-step resistive DAC, which is often used for the two-step flash subranging ADC. It should be noted how the coarse resistive ladder, in the left part of the picture, has a complete switch network, programmed by the coarse encoder, by means of which the floating fine ladder, in the right side of the picture, can be connected in parallel to any coarse resistor.

1.3.9 Interleaving, pipelining

After a brief description of the full flash and two-step flash ADCs, and before concluding the overview on the fastest data converter architectures with the pipeline the folding ADC, some considerations should be carried out about very fast ADCs. When looking into the design of a very fast ADC, a first situation that may happen is that, due to circuit limitations, some building blocks are not capable of reaching the same speed as others which, by construction or better design achievements, operate faster. Figure 1.24 illustrates this situation, where a two-step flash ADC type has a coarse flash ADC suitable for a 26MHz throughput, whilst the remaining blocks, namely S/H and fine flash ADC, do not reach this speed, but operate correctly at 13MHz.

One real explanation of the fact the coarse flash and the fine flash of Fig. 1.24, which may be even identical in the structure, do not reach the same speed, can be for example because, due to the subranging architecture, the comparators of the fine flash operate with smaller input signals, with a consequent reduction of speed. Interleaving is a technique that permits to reach the speed of the fastest building block(s), by doubling (or tripling) the hardware of the slowest structures. In Fig. 1.24, the input signal is sampled-and-held at 26MHz, by using in interleaved fashion two S/Hs, each running at 13MHz. At the output of the coarse flash ADC, the first MSB converted data are available at 26MHz, but the fine flash ADC is not capable to process this rate. In a similar manner, two fine flash ADCs are then interleaved in order to maintain an overall throughput of 26 MHz. However, the conversion rate of a classical two-step flash ADC would remain 13Msps, only one-half of the clock frequency, because the fine quantizer must wait for the result of the coarse quantizer, before executing the conversion.

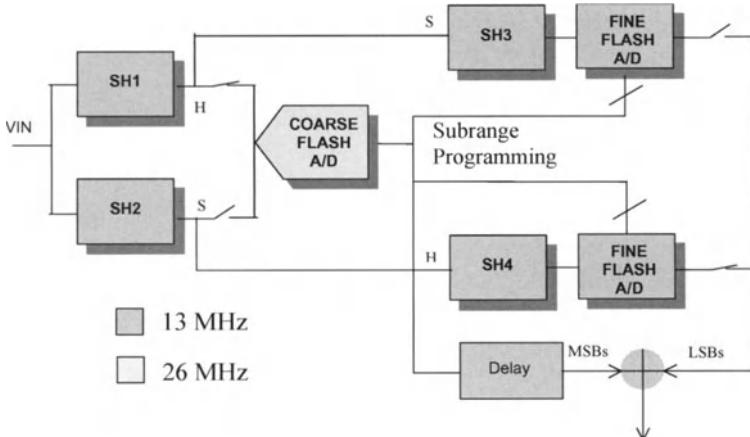


Fig. 1.24: A possible architecture for a two-step flash subranging ADC pipelined, where two input S/Hs and two fine flash ADCs are interleaved to obtain the maximum throughput of 26Msps.

In a few words, conversion speed is lost because not all blocks are “busy” all the time. Pipelining is another technique tending to increase the speed of the converter, by keeping all the blocks busy, all the time [13]. In Fig. 1.24, the coarse and the fine quantizer are pipelined by the insertion of the two S/Hs, SH3 and SH4. The input signal, after being sampled at 26MHz by the first two interleaved S/H, SH1 and SH2, is “re-transmitted”, still at a rate of 26MHz, to the fine flash ADC with a delay of a clock cycle, owing to a re-sampling carried out by SH3, SH4. The coarse and fine flash ADCs can work now in parallel at the same rate of 26MHz, and the overall converter achieves an effective sampling frequency of 26Msps. The coarse flash ADC and fine flash ADC data are recombined at the output, by inserting one clock delay to the coarse ADC output, to compensate for the systematic delay of one clock cycle the fine flash experiences with respect to the coarse flash.

1.3.10 The pipeline ADC

The exploitation of the pipelining concept, introduced in the previous paragraph as a technique to reach the maximum throughput in a data converter, leads to the pipeline ADC architecture (Fig. 1.25).

The pipeline ADC is gaining momentum among the high-speed converter architectures, especially in telecom applications, and will be largely described in the dedicated chapter of the book. It has an attractive architecture, since it is regular (typically N-cascaded stages for N-bit resolution), like the flash ADC, can still perform a conversion in one or two clock cycles, and it does not present the flash ADC design drawbacks. If we consider a multi-step flash pipelined architecture, like the one presented in the previous paragraph (Fig. 1.24), with

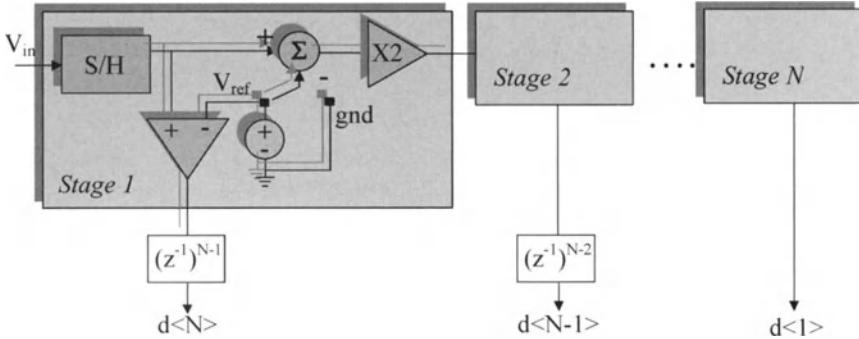


Fig. 1.25: Pipeline ADC architecture

the number of steps tending to N , and each flash quantizer with a resolution in the order of 1-bit, we obtain the pipeline ADC architecture (Fig. 1.25). At each stage, the signal is sampled-and-held, compared to a reference, hereafter i) if higher than the reference, the reference is subtracted, and an output one is produced, otherwise ii) the reference is not subtracted and an output zero is produced; finally, it is amplified by two and transmitted to the subsequent stage.

The hardware content of one stage of the pipeline, is identical to the one that was presented for the algorithmic ADC (Fig. 1.20). While the algorithmic ADC is slower by a factor of $m \cdot N$, because the unique available stage has to be used N times cyclically before converting a sample, the pipeline ADC increases the throughput by a factor N , because it can use N stages in parallel, no matter if they are delayed among them.

It can be observed that the input signal is sample-and-held N times and subject to N clock delayed bit-by-bit conversions, before getting the final digital output. This produces a so-called “latency” of N -clock cycles before the sample that enters in the pipe, is converted at the output. The N -clock cycles latency can be a drawback of the architecture in some specific applications, like control loops, where the delay can play a dominant role on the stability of the system.

1.3.11 The folding ADC

We conclude this review on Nyquist-rate converters by presenting the folding ADC architecture. The folding ADC still applies the principle of conversion of an analog signal in two steps, like the two-step flash ADC, but by paralleling them it reaches the complete conversion in one clock cycle, like the full flash ADC.

Fig. 1.26 depicts the block diagram of a 6-bit folding ADC. As for the two-step flash, the objective of the folding ADC is to dramatically reduce the hardware amount required by the full flash ADC, by keeping at the same time, the same speed target. The signal V_{in} is processed, without need of S/H, by two separate channels working in parallel. In the first one, a 3-bit flash ADC is used, while in the second an analog preprocessor, the folder, precedes the fine flash

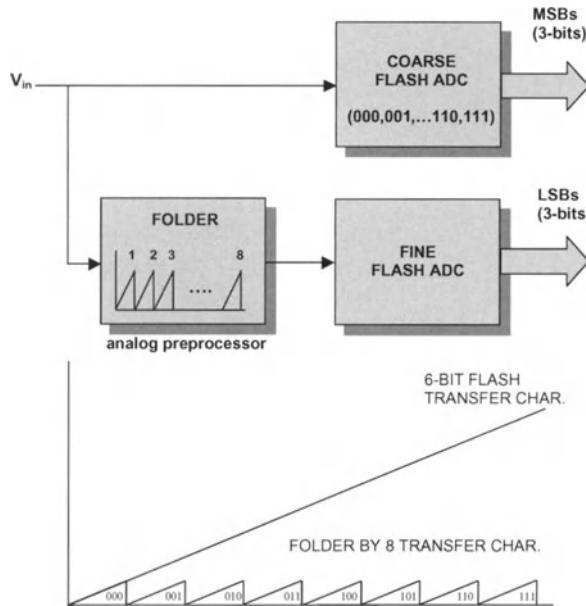


Fig. 1.26: Block diagram of a 6-bit folding ADC, with 3-bit coarse flash ADC and folder with 8-times-folding factor

ADC. The folder transfer characteristic is subdivided into $2^3=8$ identical regions, and the folder generates, into each one of them, the remaining 8 fine analog quantization levels, corresponding to the 3 LSB to be converted by the fine flash (Fig. 1.26). It turns out that for eight different input levels, a single output level corresponds, but this indetermination is solved by the flash ADC result, which identifies the unique region where the signal lies.

The folding ADC is a powerful architecture permitting to reach very high conversion rates, but the circuit implementation of the folder itself, which is key block of the ADC, can suffer of speed and accuracy limitations. Several techniques have been found in order to increase the speed of the folder, mainly based on current mode signal processing, and improve resolution and linearity, very often by means of interpolation techniques. Thanks to these improvements, the folding ADC, that was well established in bipolar technologies, has gained momentum in CMOS technology, where several implementations, achieving resolutions in the 8-bit range and conversion speeds in the 100Msps range, have been successfully demonstrated [14,15,16].

1.4 FUTURE DIRECTIONS AND CONCLUSIONS

Data converters are key technology enablers for wireless and wireline (also defined as “non-wireless”) telecom applications. These applications are pushing to the limit the electrical specifications and the technology for data converters.

Nowadays, we observe that in wireless communications, a transition by one-half generation implies roughly an increase by a factor of 10 of the data rates: from the tens kb/s peculiar of the second-generation (2G), with GSM still the king of standards, we move to the hundred kb/s of 2.5G with GPRS, up to the Mb/s in 3G with WCDMA. New emerging wireless communications beyond 3G, as the wireless LANs at 5GHz, are targeting nominal bit rates of tens Mb/s, up to the hundreds Mb/s promised by the OFDM (Orthogonal Frequency Division Multiplexing) technology. In this scenario, it appears that there will be an even more increasing demand for very high sample rates data converters. The proliferation of different standards on one side, and the need to have terminals which can ensure not only operability with, but also interoperability among, different standards (Bluetooth, GSM, WCDMA,...) on the other side, calls for data converter architectures which can be re-configurable and capable to handle very different resolutions and sample rates. Power consumption is of course another issue, especially for those circuits that are battery operated. Depending on the different type of service (voice, data, video), it is also required that the power consumption should be reduced to the lowest possible levels.

This further requirement of consumption-versus-performance optimization [17] leads to the concept of a complete programmable data conversion system, capable to deliver different sample rates and resolutions, and adapt its dissipation consequently. Another consideration is that data converters are just part of complete receive-transmit demodulation-modulation channels, and that the same considerations of re-configurability and operability with different air standards apply to these ones. In spite of this particular problem, the direct conversion architecture is gaining momentum since, due to its elementary structure, is more suitable to overcome the inflexibility of other possible choices (heterodyne, low IF,...). However, the direct conversion architecture suffers from some non-ideal effects, which often require a significant increase in the specification of the data converter in terms of resolution and effective number of bits.

In conclusion, the current and future market of telecom data converters will put more and more emphasis on those converters capable to operate at both high speed and medium-high resolutions. The design of such converters in the 8/14-bit, tens/hundreds MS/s range, is a very challenging task, but now feasible with the most advanced CMOS technologies, and the pipeline, multi-step flash, and folding data converters appear to be the best candidates to fulfill these specifications.

ACKNOWLEDGMENTS

The author would like to thank Peter Bishop of Atmel for suggestions and encouragement to write this chapter.

REFERENCES

- [1] B. H. Leung, R. Neff, P. R. Gray, R. W. Brodersen, "Area-efficient multichannel oversampled PCM voice-band coder", IEEE JSSC, vol. 23, No. 6, December 1988, pp.1351-1357.
- [2] V. Friedman, D. M. Brinthaup, D. P. Chen, T. W. Deppa, J. P. Elward, E. M. Fields, J. W. Scott, T. R. Viswanathan, "A dual-channel voice-band PCM codec using S? modulation technique", IEEE JSSC, vol. 24, No. 2, April 1989, pp. 311-317.
- [3] G. Nicollini, S. Pernici, P. Confalonieri, C. Crippa, A. Nagari, S. Mariani, A. Calloni, M. Moioli, C. Dallavalle, "A high-performance analog frontend 14-bit codec for 2.7V digital cellular phones", IEEE JSSC, Vol. 33, No. 8, August 1998, pp. 1158-1167.
- [4] F. Carassa, "Comunicazioni elettriche", (tr. Electrical communications) Boringhieri, Italy, 1983, pp. 101-111
- [5] L. Calandrino, G. Immovilli, "Sistemi di modulazione per trasmissioni numeriche" (tr. Modulation systems for numerical transmissions), Patron, Italy, 1983, pp. 26-34
- [6] A. Hairapetian, "An 81-MHz IF receiver in CMOS", IEEE JSSC Vol. 31, No. 12, December 1996, pp. 1981-1986
- [7] J. Robert, G. C. Themes, V. Valencic, R. Dessoulavy, P. Deval, "A 16-bit low voltage CMOS A/D converter", IEEE JSSC, Vol 22, April 1987, pp. 157-163
- [8] J. Robert, P. Deval, "A second-order high resolution incremental A/D converter with offset and charge injection compensation", IEEE JSSC, Vol. 23, June 1988, pp. 736-741
- [9] R. Harjani, T. A. Lee, "FRC: a method for extending the resolution of Nyquist rate converters using oversampling", IEEE Transactions on Circuits and Systems – II, Vol. 45, No.4, April 1998, pp. 482-494
- [10] J. Yuan, C. Svensson, "A 10-bit 5-MS/s Successive Approximation ADC cell used in a 70-MS/s ADC array in 1.2-um CMOS", IEEE JSSC, Vol. 29, No. 8, August 1994, pp.866-872
- [11] S. Y. Chin, C. Y. Wu, "A CMOS ratio independent and gain insensitive algorithmic analog-to-digital converter", IEEE JSSC, Vol. 31, No. 8, August 1996, pp.1201-1207.
- [12] H. S. Lee, "A 12-bit 600ks/s digitally self-calibrated pipelined algorithmic ADC", IEEE JSSC, Vol. 29, No. 4, April 1994, pp. 509-515
- [13] M. Yotsuyanagi, T. Etoh, K. Hirata, "A 10-bit 50-MHz pipelined CMOS A/D converter with S/H", IEEE JSSC, Vol. 28, No. 3, March 1993, pp. 292-300
- [14] B. Nauta, A. G. W. Venes, "A 70-MS/s 110-mW 8-b CMOS folding and interpolating A/D converter", IEEE JSSC, Vol. 30, No.12, December 1995, pp. 1302-1308
- [15] M. P. Flynn, D. J. Allstot, "CMOS folding A/D converters with current mode interpolation", IEEE JSSC, Vol. 31, No. 9, September 1996, pp. 1248-1257

- [16] A. G. W. Venes, R. J. van de Plassche, “An 80-MHz, 80-mW, 8-b CMOS folding A/D converter with distributed track-and-hold preprocessing”, IEEE JSSC, Vol. 31, No. 12, December 1996, pp. 1846-1853
- [17] R. Rivoir, F. Maloberti, G. Torelli, “Digital to analog converter with dual resistor string”, US Patent No. 08/730,592

Chapter 2

Sigma-Delta CMOS ADCs: An Overview of the State-of-the-Art

Angel Rodríguez-Vázquez, Fernando Medeiro, José M. de la Rosa, Rocío del Río, Ramón Tortosa and Belén Pérez-Verdú

2.1 INTRODUCTION^{†1}

As stated in Chapter 1, analog-to-digital conversion involves a number of tasks, namely:

- *Sampling* the input signal at frequency f_S , with prior anti-aliasing filtering and, in some cases, *holding* the sampled values.
- *Quantizing* the input sample values with N bits; i.e., mapping each continuous-valued input sample onto the closest discrete-valued level out of the $(2^N - 1)$ discrete levels covering the input signal variation interval.
- *Encoding* the result in a digital representation.

These operations were explained in Chapter 1 and are conceptually depicted in the block diagram of Fig. 2.1 for a *lowpass* signal – one whose spectrum is centered around DC. A similar diagram can be drawn for *bandpass* signals ^{†2}.

Sampling and quantization pose limitations even if realized with ideal circuit

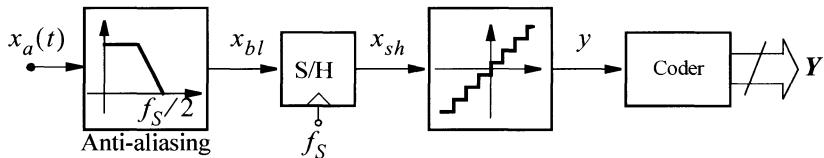


Figure 2.1: Basic operations involved in analog-to-digital conversion.

-
1. **Acknowledgement:** This work has been partially funded by the spanish MCyT and the ERDF (Project TIC2001-0929-ADAVERE), the EC IST Project 29261 (MIXMOD-EST), and the EC-IST Project 2001-34283 (TAMES).
 2. Sigma-delta converters intended for lowpass signals are called *lowpass converters* – LPΣΔCs; those for bandpass signals are called *bandpass converters* – BPΣΔCs.

components. On the one hand, sampling itself limits the frequency of the input signal ^{†3}. On the other, quantization itself corrupts the signal – the smaller the number of bits N , the larger the error. The *quantization error* is inherent in the process of mapping a continuous-valued input signal onto a set of discrete levels. Its impact is often measured through the *total in-band power of the quantization error*, i.e., the power of the error caused by the process of quantization within the signal band ^{†4}.

It is common place employing the parameter *resolution* to quantify the accuracy of the quantized signal; i.e., its similarity to the original one. Ideal circuit components would allow us to achieve any resolution for a given signal bandwidth, and vice versa. However, in practice, where sampling and quantization are realized using actual converter architectures and circuit components, a *trade-off* between bandwidth and resolution arises. This trade-off can be imagined as a negative-slope straight line in the resolution versus bandwidth plane, so that the larger the signal bandwidth, the lower the attainable resolution. It is illustrated in Fig. 2.2 where data reported for actual CMOS data converter ICs in open literature are placed in the resolution vs. conversion rate ^{†5} plane. Note that there is a particular type of converter architecture that best fits the line at each different region of the plane, the reason being that the above trade-off becomes relaxed for that architecture at the conversion-rate of the corresponding region.

Fig. 2.2 shows that Sigma-Delta Converters ($\Sigma\Delta C$) cover a very wide region of this plane. During the 80's $\Sigma\Delta C$ IC-implementations were basically devised for audio [Adam86] [Bose88]. Over the years, they have been also employed to replace incremental and integrating architectures (see Section 1.3 in Chapter 1)

-
3. For lowpass signals this limitation affects only the *bandwidth* B_W of the signal. For bandpass signals, it affects the *bandwidth* B_W and the *location* f_n of the band.
 4. Under the assumptions in Chapter 1 and considering an ideal quantizer with uniform quantization, the in-band power of quantization error is calculated as,

$$P_{eq} \approx \frac{\Delta^2}{6} \cdot \frac{B_W}{f_S} = \frac{\Delta^2}{12} \cdot \frac{2 \cdot B_W}{f_S}$$

where $\Delta = X_{FS}/(2^N - 1)$ represents the amplitude of each quantization interval, called V_{LSB} in Chapter 1, and X_{FS} is the full-scale range of the quantizer. This expression assumes that the quantizer does not *overload*; i.e. that the input is such that the modulus of the difference between input and output is smaller than $\Delta/2$. For inputs exceeding the full-scale quantizer range, the output saturates, and the difference between input and output, as well as the power of error, increase.

5. The conversion rate is the Nyquist frequency; i.e., for lowpass signals, twice the signal bandwidth.

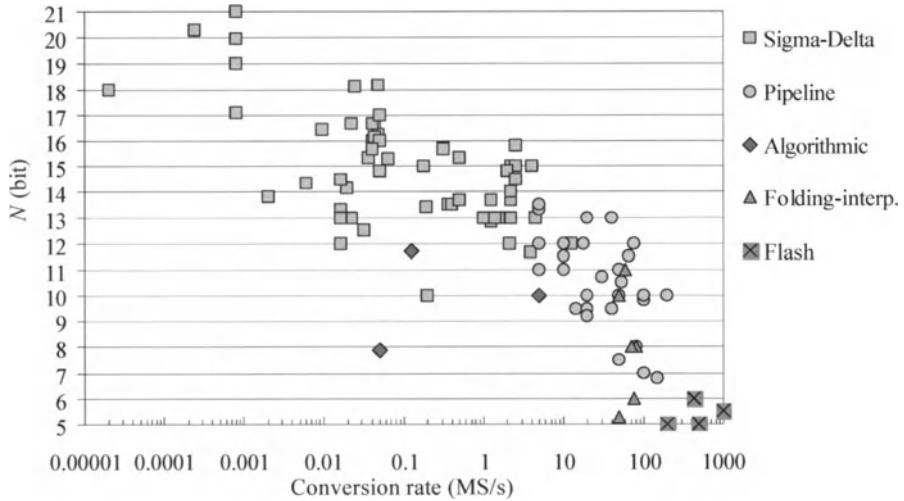


Figure 2.2: CMOS ADC ICs in the resolution vs conversion rate plane.

for sensor applications [Bertl93] [Briga02] [Kert94] [Nys97] [Yama94], and today they are being extensively used for wireline [Casi01] and wireless [Galt02] communication systems. Such a pervasive usage is partly due to the potential ^{†6} advantages of $\Sigma\Delta$ Cs regarding linearity and insensitivity to analog inaccuracies. These advantages are the consequence of two features that are inherent in the operation of $\Sigma\Delta$ Cs, namely:

- They exchange bandwidth for accuracy; to this purpose, more signal samples than needed are used to implement *time averaging*, thus increasing resolution. On the contrary, full Nyquist converters (see Section 1.3 in Chapter 1) sample at the minimum possible rate, defined by the *Nyquist frequency*,

$$f_N = \begin{cases} 2 \cdot B_W & , \text{ for low-pass signals} \\ \frac{2 \cdot (f_n + B_W/2)}{\left(f_n + \frac{B_W}{2}\right)/B_W} & , \text{ for band-pass signals} \end{cases} \quad (2.1)$$

$\Sigma\Delta$ Cs sample at a rate M times higher than the minimum required,

$$f_S = M \times f_N \quad (2.2)$$

where M is called the *oversampling ratio* ^{†7}.

-
6. As operation frequency increases, the potential advantages of sigma-delta architectures become less evident, and involved analog design is needed for performance.

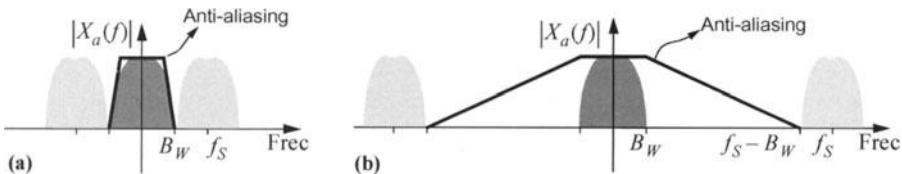


Figure 2.3: Illustrating the anti-aliasing filter attenuation for (a) full-Nyquist sampling; (b) oversampling.

The attenuation is smaller when oversampling is used. Even for moderate oversampling, such as $M = 16$, the attenuation is low enough to enable using simple anti-aliasing filters. For instance, a passive second-order filter suffices for an ADSL $\Sigma\Delta$ Cs [Rio02b].

Oversampling has two noticeable effects. On the one hand, it yields larger separation of the images of the input signal spectrum created during the sampling process. Consequently, the selectivity of the filter used to avoid aliasing of these images, the *anti-aliasing* filter, is less demanding than for full Nyquist converters – see Fig. 2.3. On the other hand, when an oversampled signal is quantized, the spectral components of the quantization error are distributed in a larger frequency range. Consequently, the *Power Spectral Density (PSD)* of the error decreases within the signal band – see Fig. 2.4. Obviously, in order to take advantage of this feature we must be capable of removing the out-of-band error components without degrading the *baseband*

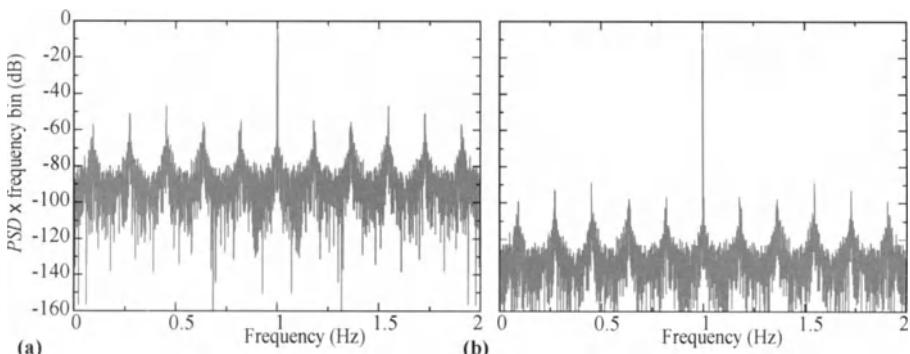


Figure 2.4: *PSD* of two oversampled 1V, 1Hz sinusoid quantized with $N = 3$:
(a) $M = 8$; (b) $M = 1024$.

Note that the noise floor within the band is larger for $M = 8$.

-
7. The concept of sigma-delta conversion can be used with no oversampling, i.e. with $M = 1$. However, in this chapter we implicitly assume that oversampling does exist, i.e. that $M > 1$. Hence, rigorously speaking we should refer to the converters in this chapter as *oversampled sigma-delta converters*. However, for simplicity we simply call them sigma-delta converters.

(or signal band); an operation that requires high-selectivity filters. It may be argued that this poses a drawback similar to the necessity of high-selectivity anti-aliasing filters in full Nyquist converters. However, while anti-aliasing filters handle analog signals, the filters employed to remove the out-of-band components of the quantization error handle digital signals, and designing high-selectivity digital filters is much easier than designing high-selectivity analog ones.

- Another inherent feature in $\Sigma\Delta$ Cs is the use of *feedback* to attenuate^{†8} the errors introduced in the quantization process. As in any other feedback systems, in $\Sigma\Delta$ Cs the encoded output signal is “compared” to the input and the “difference” is forced to be zero by the action of feedback. Fig. 2.5 illustrate the input and output (feedback) signals for the case in which the latter is quantized with a single bit (only two quantization levels, ± 1), giving rise to a quickly-varying (oversampled) *pulse stream* or *bit stream*. In this example the input signal sweeps five different levels. Note that the pulse density, i.e. the relative number of $+1$ and -1 ’s, is different for each level, so that the averaged output coincides with the input. For zero input level, the number of positive and negative pulses is practically the same, thus nullifying the average of the feedback signal. For positive (alternatively negative) inputs, such average turns positive (alternatively negative), always tracking the input level.

$\Sigma\Delta$ Cs combine feedback and oversampling to reproduce the input signal

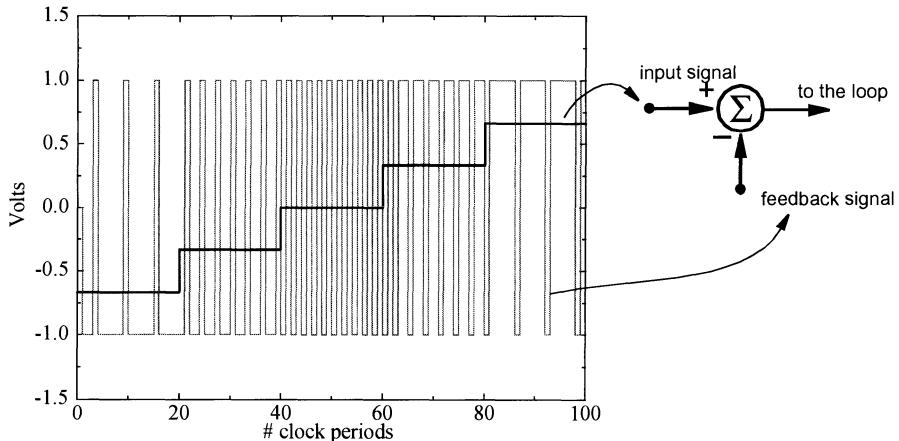


Figure 2.5: Pulse stream for encoding an input signal sweeping several levels as obtained by using a first-order sigma-delta converter.

-
8. Attenuation happens only within the signal band. For frequencies outside this band errors are actually enlarged.

with the largest possible fidelity via the in-band components of the pulse stream, while placing the majority of the error power outside the signal band. As in other feedback systems, this can be done very efficiently, but there is a snag: feedback dynamics may produce *instability*. In $\Sigma\Delta$ Cs such instability can generate rather involved, exotic behaviors because strongly nonlinear blocks (quantizers) are inserted in feedback loops with sometimes high-order dynamics.

Features underlying the operation of sigma-delta modulators render them a good choice to realize embedded analog-to-digital interfaces in modern system-on-chip ICs. The theory of sigma-delta is already very well covered by a number of books (see for instance [Enge99] [Mede99a] [Nors97] [OpT93] [Rosa02] and the many references listed in these works). This chapter is not aimed at replacing these books. Also, the contents of the chapter have been selected trying to avoid duplication with other chapters in this book. Although some duplication is unavoidable, the basic purpose here is providing an overview of the current state-of-the-art on sigma-delta CMOS implementations, without entering into detailed descriptions of architectures and circuits. These solutions, particularly those adopted for telecom applications, are found in Chapters 8, 9, 10, and 11.

2.2 ARCHITECTURE AND COMPONENTS OF $\Sigma\Delta$ Cs

Because $\Sigma\Delta$ Cs are feedback systems, they do not fit well into the conceptual block diagram in Fig. 2.1. The sampler and the quantizer of this diagram are cascaded in an open loop configuration, like actually happens in full Nyquist flash converters. Instead, in $\Sigma\Delta$ Cs the sampler and the quantizer are connected, together with other dynamic blocks, into a feedback structure called *Sigma-Delta Modulator* ($\Sigma\Delta$ M). Hence, the diagram of a $\Sigma\Delta$ C is better represented through Fig. 2.6, where Fig. 2.6(a) corresponds to lowpass converters and Fig. 2.6(b) is for band-pass.

The diagrams of Fig. 2.6 comprise three large blocks, whose operation is briefly described below and illustrated in Fig. 2.7 for the lowpass case:

- 1) *Anti-Aliasing Filter*. The function of this block should be already clear.
- 2) *Sigma-Delta Modulator*. It takes the filtered analog signal, samples and quantizes it by using a feedback system which embeds dynamic sub-blocks, and delivers a stream of digital words at the sampling rate; each sample of this stream encodes the quantized value of the concurrent quantizer input sample. Unlike flash quantizers, where the quantizer input coincides with the converter input, in sigma-delta modulators both inputs are different.

Since the sampling rate is higher than needed due to oversampling, the

stream obtained at the modulator output is a *high-frequency* one. However, the information of interest is contained only at low-frequencies; in the *baseband*, i.e within the signal bandwidth. There, and due to the feedback action, the input signal is reproduced with much larger accuracy than featured by the embedded quantizer itself. Actually, dynamic feedback action *shapes* the quantizer errors in frequency domain, and *pushes* a significant part of these errors towards high-frequencies, outside the signal band. This is the reason why sigma-delta modulators are also called *noise-shaping* modulators, where the term noise refer to the errors introduced during the process of quantization, as explained in Chapter 1.

- 3) *Error-Removal Filter + Decimator*. This purely digital block takes the high-frequency stream as input, removes its high-frequency components through high-selectivity digital filtering, and decimates it to reduce the sampling frequency down to the minimum required, i.e. the Nyquist frequency. The result is the input signal converted at the Nyquist rate with a resolution that is much larger than that of the quantizer embedded in the modulator.

Note that the quantizer embedded in either Fig. 2.6(a) or Fig. 2.6(b) has B bits. Hence, when operated at the Nyquist rate it produces, by itself, an error whose power within the signal band is given by (see footnote 4),

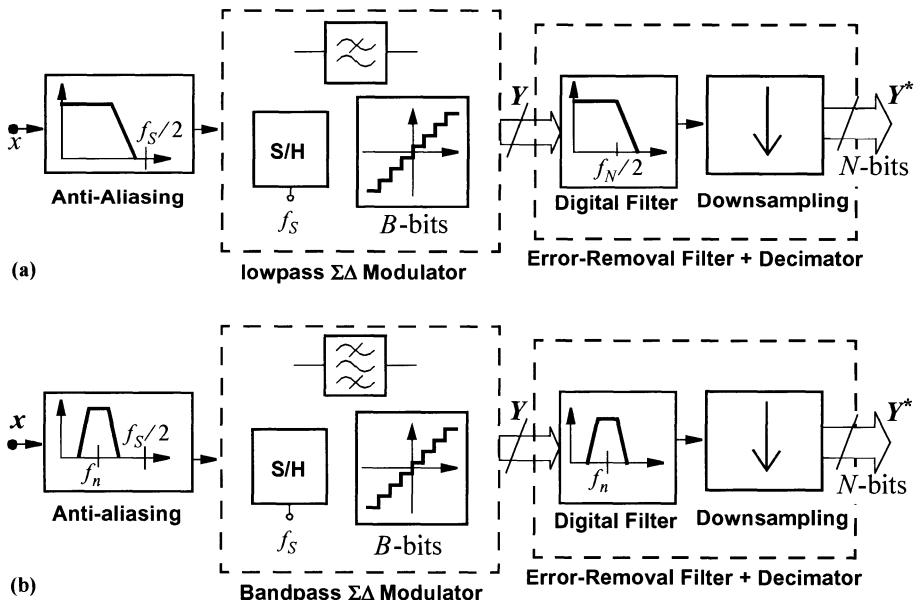


Figure 2.6: Block diagram of a sigma-delta converter: (a) LP $\Sigma\Delta$ C; (b) BP $\Sigma\Delta$ C.

$$P_{eq} \approx \frac{{X_{FS}}^2}{12 \cdot (2^B - 1)^2} \quad (2.3)$$

assuming that there is no *overloading*^{†9}. However, when the quantizer is embedded in a sigma-delta loop, the in-band error power of the stream at the decimator output happens to be,

$$P_{eq} \approx \frac{{X_{FS}}^2}{12 \cdot (2^N - 1)^2} \quad (2.4)$$

where $N = \alpha \times B$, and α typically ranges from 5 to 18. It means that the in-band

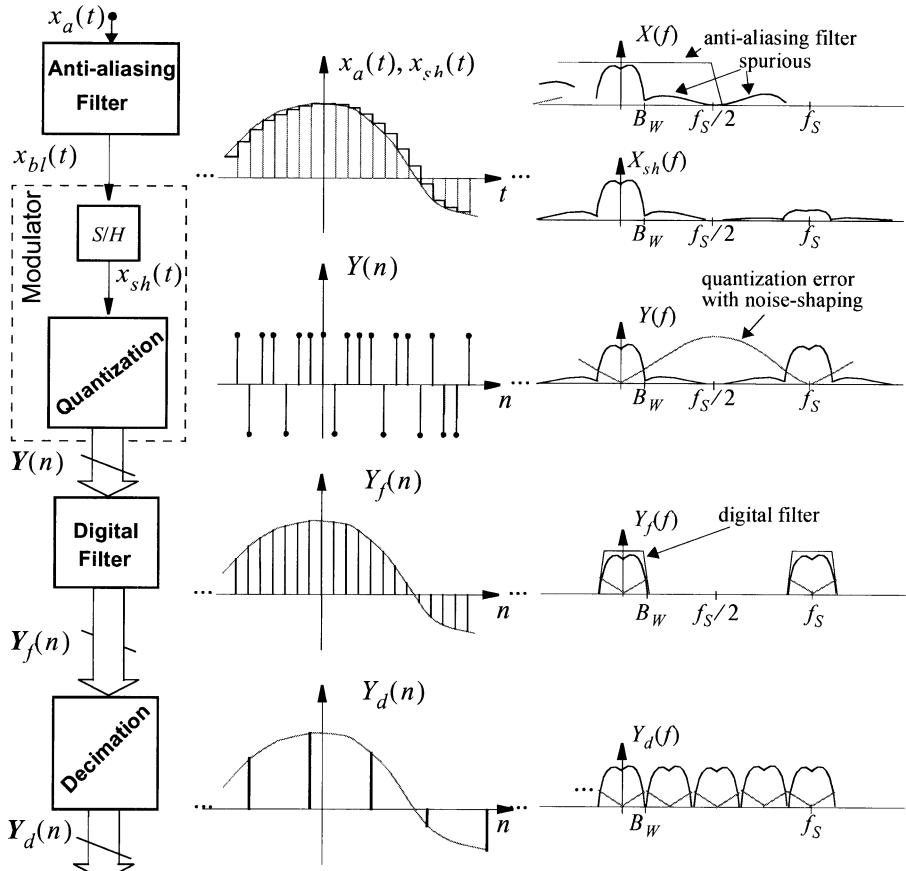


Figure 2.7: Illustrating the signal processing operation flow of a LPΣΔCs.

The quantizer embedded in the modulator is assumed to be of just one bit ($B = 1$).

(baseband) error at the output corresponds to the value that would be obtained by using a quantizer with a number of bits, N , larger than B ; i.e. larger than the number of bits of the quantizer embedded in the sigma-delta modulator itself.

The basic responsible for this reduction of the in-band error power, and associated resolution enhancement, is the $\Sigma\Delta$ modulator. Hence, the design of this block creates significant challenges to effectively achieve the pursued accurate encoding of the in-band signal components, while making the circuitry operate at high-frequencies. From now on, we will hence focus on this block. However, designers should not forget that a $\Sigma\Delta$ converter is more than just a $\Sigma\Delta M$.

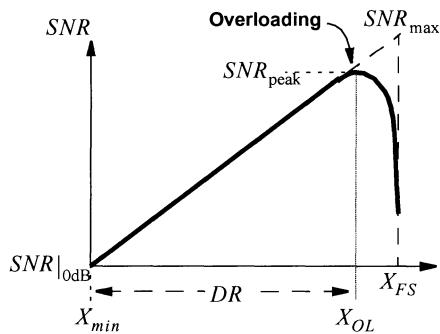
2.3 INGREDIENTS OF SIGMA-DELTA MODULATORS

Refer to Fig. 2.8(a) and consider first an isolated B -bit quantizer sampled at f_S . Assume that each sample of the quantizer output is coded by a digital word Y . From this digital output, the analog quantized version y of the input signal is obtained by using a Digital-to-Analog Converter (DAC); i.e. by assigning an analog value (*reference voltage*) to the full-scale of the digital code, and mapping code values according to the chosen quantization law (see Chapter 1) ^{†10}. Obviously, if we subtract x from y the outcome is not null. We already know that the process of quantization yields an error,

$$e_q = y - x \quad (2.5)$$

even if accomplished in fully ideal way. Each sample of this error signal has a

9. Assuming a sinusoidal input, the maximum peak-to-peak amplitude which does not produce overloading is $X_{FS} + \eta \cdot \Delta \approx X_{FS}$, where η is close to unity. When the quantizer is placed within a $\Sigma\Delta M$, overloading happens at smaller amplitudes. In either case, beyond overloading the error power increases with the signal, thus making the SNR to decrease. It is illustrated in the figure right where the smaller the overloading limit X_{OL} , the lower



SNR_{peak} . Increasing the overloading limit is an important design concern indeed.

10. For instance, if the quantizer is 1-bit, the digital signal consists of a stream of “ones” and/or “zeroes”; while the analog signal is a stream of pulses with amplitude $X_{FS}/2$ and/or $-X_{FS}/2$. Usually E_r (voltage reference) is used to denote $X_{FS}/2$.

value equal to the difference between the concurrent input and the corresponding quantization level (see Fig. 1.3 in Chapter1). If the quantizer does not overload, the amplitude of this error remains smaller than $\Delta/2$ at any time, with Δ being the separation between consecutive levels of the DAC output (see footnote 4).

It is important to be aware that the quantization error is completely dependent on the input signal. However, as it was already explained in Chapter 1, under certain assumptions the quantization error can be handled as an additive white noise contribution which is not correlated to the input and whose power is *uniformly distributed* in the band $[-f_S/2, f_S/2]$, as depicted in Fig. 2.8(b) ^{†11}.

2.3.1 Oversampling

This is the first ingredient of $\Sigma\Delta$ Ms as considered in this Chapter (see footnote 7). The benefit of oversampling is evident from Fig. 2.8(b), which corresponds to $M = 8$. Since the error power is distributed over a band larger than the signal band, only a portion of the error happens to be in-band; the larger M the smaller the portion. This portion is given by,

$$P_{eq} \approx \frac{\Delta^2}{12} \cdot \frac{1}{M} \quad (2.6)$$

By taking into account that $\Delta^2 = X_{FS}^2 / (2^B - 1)^2$, combining the equation above

11. Be aware that this is just an assumption. The quantization error is entirely dependent on the input signal, and hence it is not an additive noise. However, as demonstrated in:

W. Bennett, "Spectra of Quantized Signals". *Bell Syst. Tech. J.*, Vol. 27, pp. 446-472, July 1948.

B. Widrow, "A Study of Rough Amplitude Quantization by Means of Nyquist Sampling Theory". *IRE Trans. on Circuit Theory*, Vol.3, pp. 266-276, Dec. 1956.

A.B. Sripad and D.L. Snyder, "A Necessary and Sufficient Condition for Quantization Errors to be Uniform and White". *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 25, pp. 442-448, Oct. 1977.

M.R. Gray, "Quantization Noise Spectra". *IEEE Trans. on Information Theory*, Vol. 36, pp. 1220-1244, November 1990.

under certain conditions the errors caused by quantization can be modelled as an independent additive white noise source. The conditions are: a) the quantizer does not overload (the error amplitude remains smaller than $\Delta/2$; b) the input signal is a random signal; c) the characteristic function of the input signal (the Fourier transform of its probability density function) is band limited. Although these conditions are hardly met in practice, the additive noise assumption is common place in $\Sigma\Delta$ Cs design – the smaller Δ (i.e. the larger the number of bits in the quantizer) the better the assumption. Whenever needed, a dither signal can be added to modify the statistical properties of the quantizer input signal and thus the validity of the additive noise model:

S.P. Lipshitz, R.A. Wannamaker, and J. Vanderkoy, "Quantization and Dither: A Theoretical Survey". *J. Audio Eng. Society*, Vol. 40, pp. 355-375, May 1992.

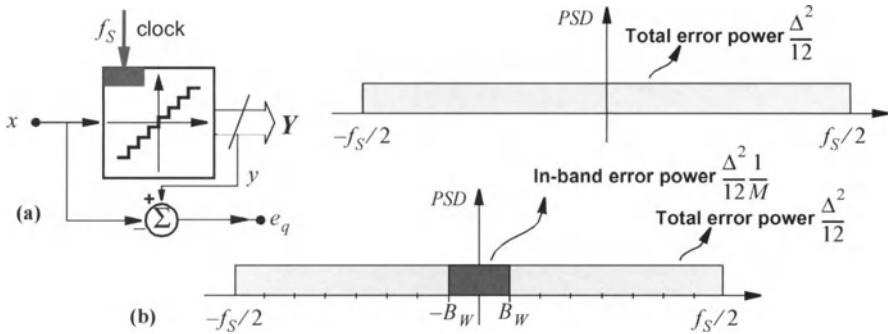


Figure 2.8: Illustrating the benefits of oversampling: (a) Sampled quantizer symbol and approximated PSD of the quantization error; (b) In-band error power (highlighted) and total error power.

The DAC needed to obtain y from Y is not explicitly shown.

with (2.4), and assuming $2^N \gg 1$ and that M is large enough one obtains,

$$N \approx \frac{1}{6,02} \cdot [20 \cdot \log(2^B - 1) + 10 \cdot \log(M)] \quad (2.7)$$

which by defining $M = 2^\xi$ can be reformulated into,

$$N \approx \frac{20 \cdot \log(2^B - 1)}{6,02} + \frac{\xi}{2} \quad (2.8)$$

It shows that, by just oversampling, the equivalent resolution at the output increases one bit every time the oversampling ratio is quadrupled ^{†12}. Enhancement caused by oversampling was illustrated in Fig. 2.4 by showing the PSD of one quantized sinusoid for two values of the oversampling ratio.

12. The above considerations have been made by using the in-band error power. Alternatively, the maximum signal-to-noise ratio SNR_{\max} can be used. It is obtained by considering a sinusoidal input with the maximum amplitude, see (1.12) in Chapter 1, as

$$SNR_{\max} \approx 6,02 \cdot B + 1,76, \text{ dB}$$

for a quantizer with B bits sampled at the Nyquist frequency. It shows that every bit contributes approximately 6.02 dB to the SNR_{\max} . When oversampling is used, the maximum signal-to-noise ratio is recalculated as,

$$SNR_{\max} \approx 6,02 \cdot B + 1,76 + 6,02 \cdot \frac{\xi}{2}, \text{ dB}$$

where $M = 2^\xi$. It confirms that one bit (i.e. 6.02 dB) is gained every time the oversampling ratio is multiplied by a factor of 4.

2.3.2 Error Processing

Although oversampling itself enhances the equivalent number of bits, the rate of enhancement is rather slow. For instance, to obtain 8bits within 10kHz bandwidth by just oversampling a 1-bit quantizer, the sampling frequency should be around 328MHz. This is obviously a rather inefficient strategy – clever methods must be devised.

Consider lowpass $\Sigma\Delta$ Ms. Bear in mind that the target is reducing the quantization errors at low frequencies; i.e. within the signal band. For large enough oversampling ratios, with the signal hardly changing from sample to sample, most of the changes in the quantization error happen at high frequencies; i.e. low frequencies components of consecutive quantization error samples are almost identical. Hence, these low-frequency components can be attenuated by subtracting the previous sample from the current one,

$$e_{q\text{HP}}(n) = e_q(n) - e_q(n-1) \quad (2.9)$$

The efficiency of this strategy is confirmed via Fig. 2.9, showing a reduction of the power of error by a factor close to 20.

Further reduction is achieved by involving more error samples,

$$\begin{aligned} e_{q\text{HP}}(n) &= e_q(n) - e_q(n-1) && , \quad \text{1st order error proc.} \\ e_{q\text{HP}}(n) &= e_q(n) - 2 \cdot e_q(n-1) + e_q(n-2) && , \quad \text{2nd order error proc.} \\ e_{q\text{HP}}(n) &= e_q(n) - 3 \cdot e_q(n-1) + 3 \cdot e_q(n-2) - e_q(n-3) && , \quad \text{3rd order error proc.} \\ &\dots && , \quad \dots \end{aligned} \quad (2.10)$$

The procedure is formulated in an unified manner by resorting to z -transforms,

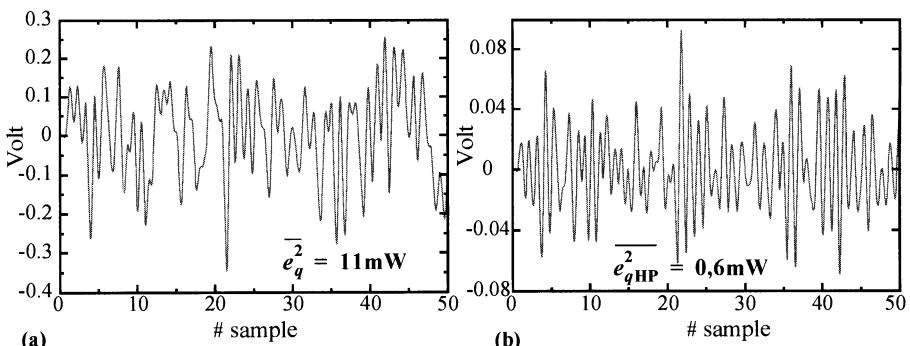


Figure 2.9: Illustrating the benefits of error processing: (a) error waveform and power without processing; (b) the same with the processing in (2.9). Input is a full-scale (1V) random signal lowpass filtered at 1Hz. Sampling frequency is 16Hz. Quantizer is one bit. Note that the scales of the vertical axis in (a) and (b) are different.

$$E_{q\text{HP}}(z) = (1 - z^{-1})^L \cdot E_q(z) \quad (2.11)$$

It shows that the processed error is a filtered version of the original, the *filtering* transfer function being,

$$N_{TF}(z) = (1 - z^{-1})^L \quad (2.12)$$

where L denotes the *order* of the filtering realized on the quantization error.

The magnitude $|N_{TF}(e^{j\Theta})|$ of this transfer function is,

$$|N_{TF}(e^{j\Theta})| = |1 - e^{-j\Theta}|^L = 2^L \cdot \sin^L\left(\frac{\Theta}{2}\right) \quad (2.13)$$

where $\Theta = 2\pi \cdot \frac{f}{f_S} = \pi \cdot \frac{1}{M} \cdot \frac{f}{B_W}$. Note that in-band is $f \leq B_W$, and $\Theta \ll 1$, if M

is large enough. Hence, the transfer function is very small within the signal band and this results in the processed error having very small in-band power,

$$P_{\epsilon q} = \int_{-B_W}^{B_W} \frac{\Delta^2}{12} \cdot \frac{1}{f_S} \cdot |N_{TF}(f)|^2 \cdot df \approx \frac{\Delta^2}{12} \cdot \frac{\pi^{2 \cdot L}}{(2 \cdot L + 1) \cdot M^{(2 \cdot L + 1)}} \quad (2.14)$$

much smaller than if only oversampling were applied – compare with (2.6).

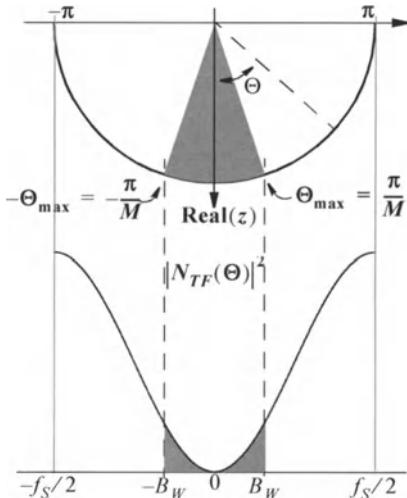


Figure 2.10: Amplitude of N_{TF} for 1st order error processing.

Fig. 2.10 depicts the squared magnitude of N_{TF} for $L = 1$ and suggests that error reduction happens due to:

- 1) the *highpass* shape of the noise transfer function, which yields very large attenuation in-band and pushes the noise towards high frequencies;
- 2) the inverse dependence of the integration interval with the oversampling ratio.

Note that the highly nonlinear nature of the shaping function results into a very nonlinear dependency of $P_{\epsilon q}$ with M – see (2.14). It means that the influence of oversampling is largely amplified, which makes manifest as a significant

resolution enhancement,

$$N \approx \frac{1}{6,02} \left[20 \cdot \log(2^B - 1) + 20 \cdot \log(M) \cdot \left(L + \frac{1}{2} \right) + 20 \cdot \log\left(\frac{\sqrt{2 \cdot L + 1}}{\pi^L}\right) \right] \quad (2.15)$$

and it is further illustrated in Fig. 2.11, where SNR_{max} is plotted as a function of the order of the filtering for several oversampling ratios and a single bit quantizer ($B = 1$).

From (2.15) and denoting $M = 2^\xi$, the resolution enhancement due to filtering plus oversampling is approximately given by,

$$\Delta N \approx \xi \cdot \left(L + \frac{1}{2} \right) \Big|_{\text{bits}} \quad (2.16)$$

meaning that $(L + 1/2)$ bits are gained every time the oversampling is doubled.

The same concept applies for bandpass $\Sigma\Delta$ Ms. In this case the error processing transfer function can be obtained by making the following transformation ^{†13},

$$z^{-1} \rightarrow -z^{-2} \quad (2.17)$$

which transforms the highpass transfer function of the LP case into a *notch* transfer function centered around $f_n = f_S/4$,

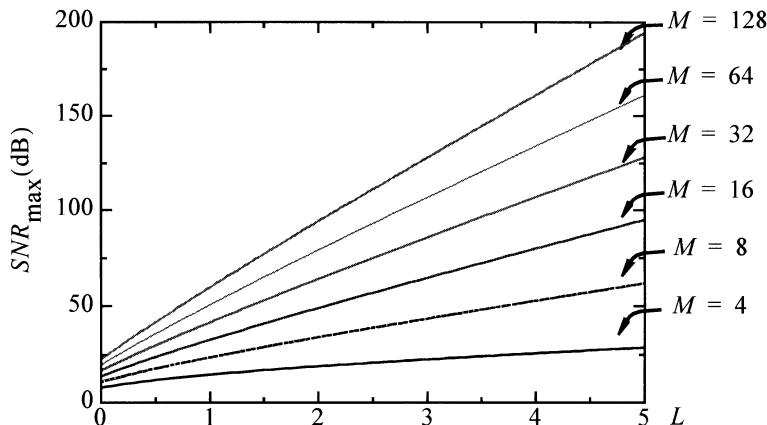


Figure 2.11: Illustrating the combined action of noise filtering and oversampling on the maximum signal-to-noise ratio.

13. This is not the only possibility as explained in Chapter 11.

$$N_{TF}(z) = (1 + z^{-2})^L \quad (2.18)$$

and results in the same resolution enhancement as (2.16).

2.3.3 Feedback

This is the third ingredient of $\Sigma\Delta$ Ms. It is needed to put the error processing strategy into practice. Without feedback, the error processing method explained above cannot be implemented. Refer to Fig. 2.8(a). The digital output Y is a mixture of the input signal and the quantization error. Processing that signal modifies both the input and the quantization error. Hence, successful implementation of the error processing method requires input and quantization error to be, first, separated and, then, separately processed. The only way to achieve this is placing the quantizer within a feedback loop^{†14}.

A first possibility to implement error processing through feedback is using the configuration in Fig. 2.12(a), which includes a *loop filter* in the feedback path. There, assuming that the quantizer is represented by the linear, additive noise model of Fig. 2.12(b) with $G_q = 1$, and that the system is stable, one obtains,

$$\left. \begin{array}{l} U(z) = X(z) - H(z) \cdot E_q(z) \\ Y(z) = U(z) + E_q(z) \end{array} \right\} \Rightarrow Y(z) = X(z) + [1 - H(z)] \cdot E_q(z) \quad (2.19)$$

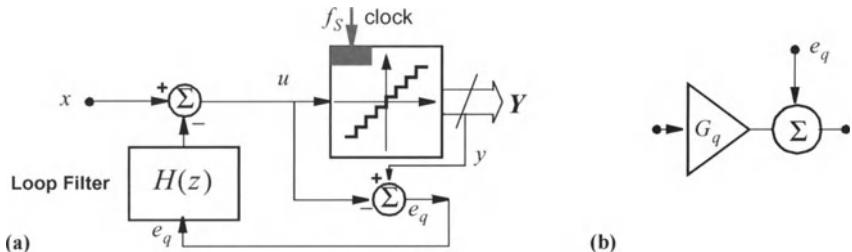


Figure 2.12: Error processing implementation through feedback.

C.C. Clutler, *Transmission System Employing Quantization*. U.S. Patent No. 2,927,962, 1960.

-
14. The quantization error produced by a first quantizer can be extracted by subtracting the input from the analog reconstructed quantizer output. Then, the quantization error can be quantized and encoded by a second quantizer, and the digital outputs of both quantizers can be combined in digital domain. Since the error of the first quantizer is present at both digital outputs, it can be annulled through proper processing in that digital domain. Thus, only the error of the second quantizer remains. This method is employed in many practical sigma-delta converter architectures and circuits. However, feedback is still needed to attenuate the second quantizer error.

The top-left equation reveals the principle. Note that the quantizer input contains the input signal x and the filtered version of the quantization error that is fed back. Hence, the outcome of the quantization process (see bottom-left equation) contains two representations of the quantization error. By properly choosing $H(z)$ the noise-shaping concept explained in Section 2.3.2 can be implemented.

Although conceptually simple, the architecture in Fig. 2.12 has two practical problems. On the one hand, it raises significant *stability*^{†15} issues for high-order error processing functions; on the other, it is very sensitive to the errors in the analog substractor employed to extract the quantization error.

These drawbacks are overcome by placing the loop filter in the feedforward path, and using the output signal (instead of the quantization error) as feedback signal – see Fig. 2.13, where the lowpass and the bandpass cases are included. The idea, which was already outlined in Fig. 2.5, is simple. The loop filter renders the loop gain frequency-dependent; inside the signal band it is very high and outside very low. As a result, and due to the action of feedback, the error signal $\hat{u} = x - y$ becomes practically null in the signal band. Consequently, x and y practically coincide within this band – the larger the loop gain the smaller the discrepancies. As

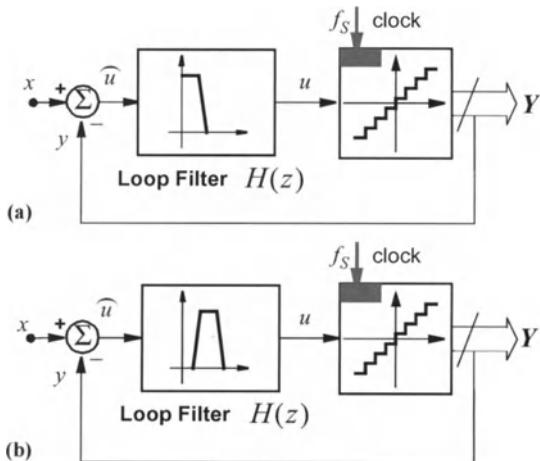


Figure 2.13: LP and BP $\Sigma\Delta M$ architectures.

Pioneering contributions include:

H. Inose, Y. Yasuda and J. Murakami, “A Telemetering System by Code Modulation- Δ - Σ Modulation”. *IRE Trans. on Space Electronics and Telemetry*, Vol. 8, pp. 204-209, September, 1962.

P.H. Gailus, W.J. Turney, and F.R. Yester, *Method and Arrangement for a Sigma Delta Converter for Bandpass Signals*. U.S. Patent 4,857,828, filed Jan. 28 1988, issued Aug. 15 1989.

15. Sigma-delta modulators are highly non-linear dynamic systems whose stability is difficult to study. Theoretically, stability means that the *state variables* of the system remain bounded for certain class of inputs. In practice, however, it must be also guaranteed that state variables take small enough values. Instability is associated to uncontrolled growing of the signal amplitudes. A modulator is stable if, for bounded inputs and whatever integrator initial conditions, the internal state variables remain also bounded over time.

a by-product, these discrepancies, which assuming ideal components are only due to quantization errors, are pushed outside the signal band where they can be removed by using high-selectivity digital filters as explained before.

The architectures in Fig. 2.13 are analyzed by using the linear, additive-noise quantizer model in Fig. 2.12(b)^{†16} to obtain,

$$Y(z) = \frac{G_q \cdot H(z)}{1 + G_q \cdot H(z)} \cdot X(z) + \frac{1}{1 + G_q \cdot H(z)} \cdot E_q(z) \quad (2.20)$$

This equation shows that the input and the quantization error are affected by different transfer functions, and hence may exhibit different frequency-dependent behaviors. On the assumptions made in preceding paragraphs:

$$S_{TF}(z) \equiv \frac{G_q \cdot H(z)}{1 + G_q \cdot H(z)} \approx \begin{cases} \equiv 1 & , \text{within the signal band} \\ \text{don't care} & , \text{outside the signal band} \end{cases} \quad (2.21)$$

$$N_{TF}(z) \equiv \frac{1}{1 + G_q \cdot H(z)} \approx \begin{cases} \ll 1 & , \text{within the signal band} \\ \gg 1 & , \text{outside the signal band} \end{cases}$$

By properly choosing the loop filter $H(z)$, noise shaping transfer functions in (2.12), for lowpass, and (2.18), for bandpass (or others yielding similar results) are built, thus implementing the noise shaping concept illustrated in Fig. 2.14.

2.4 CLASSIFICATION OF $\Sigma\Delta$ M IC ARCHITECTURES

Out of the many architectures reported in literature only those that have been realized in IC form are considered for classification here. They can be grouped attending to different criteria:

- The nature of the signals being handled: lowpass versus bandpass $\Sigma\Delta$ s.
- The type of dynamics of the loop filter. Previous sections have assumed discrete-time dynamics everywhere. $\Sigma\Delta$ s having this type of dynamics are called *Discrete-Time* (DT). There is another type of modulators, called *Continuous-Time* (CT), whose loop filter is continuous-time whereas the quantizer is discrete-time.

16. It is worth insisting that studying sigma-delta modulators using linear models represents an oversimplified approach. This approach is normally adopted to design IC modulators in the presence of circuit non-idealities. However, designers should be fully aware of the many non-linear phenomena that may go unnoticed when using these linear models. For a comprehensive coverage of these phenomena, the interested reader is referred to [Enge99].

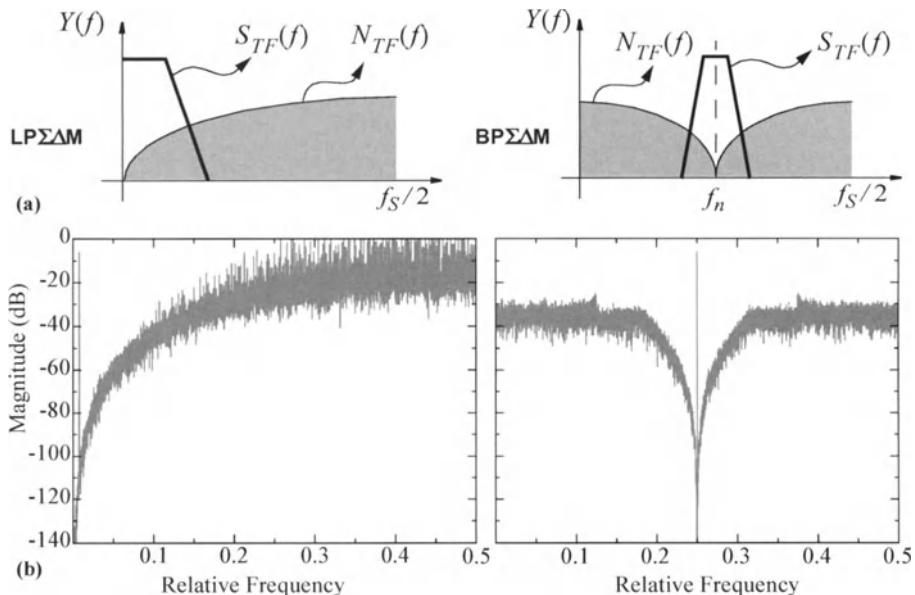


Figure 2.14: Illustrating the noise shaping concept in $\Sigma\Delta$ Ms: (a) Conceptual illustration; (b) Spectra associated to 4th-order noise-shaping modulators.

In the spectra at the bottom, the input is a sinusoid of amplitude $X_{FS}/2$ and frequency $f_S/128$ and $f_S/4$, for the lowpass and bandpass modulator, respectively.

- The number of bits in the embedded quantizer. Historically, modulators employed mostly *single-bit* quantizers ($B = 1$) because these quantizers do not raise *Integral Non-Linearity* ^{†17} (INL) issues and are very easy to design. Today, the use of *multi-bit* quantizers ($B > 1$) has been spread.
- The number of quantizers employed. $\Sigma\Delta$ Ms employing only one quantizer are called *single-loop* structures. Those employing several quantizers have different names: cascade, dual-quantization, truncation-feedback, etc.
- Type of circuitry employed, primitives available in the fabrication technology, voltage supply, etc. Most of the reported DT implementations employ switched-capacitor circuits with high-quality passive capacitors – mixed-signal technology options. Others employ passive capacitor structures available on standard CMOS technologies. There are also those that can be implemented with just transistors, such as those employing MOS devices as capacitors, or those that use switched-current circuits, etc.
- Performance featured by the modulator. Usually this is characterized by the

17. The concept of integral non-linearity has been already introduced in Chapter 1.

resolution (measured in terms of the equivalent number of bits of the output signal), the maximum *bandwidth* of the input signal, and the *power* consumption of the circuit. These features can be combined in a Figure-Of-Merit (*FOM*) to quantify the “quality” of a given circuit.

Two different *FOM*s are typically employed for LPΣΔMs, namely:

$$FOM_1 = \frac{P_w}{2^{ENOB} \times 2 \cdot B_W} \cdot 10^{12} \Bigg|_{\text{pJ/(conversion step)}} \quad (2.22)$$

proposed by Goodenough^{†18}, and,

$$FOM_2 = 2 \cdot k \cdot T \cdot \frac{3 \times 2^{2 \cdot ENOB} \times 2 \cdot B_W}{P_w} \quad (2.23)$$

proposed by Wooley^{†19}. P_w in both figures denotes the power consumption of the modulator in Watts; B_W is the modulator bandwidth in Hz; and $ENOB$, the equivalent number of bits, is related to the SNR_{\max} according to the formula $SNR_{\max} \approx 6,02 \cdot ENOB + 1,76$, (dB).

FOM_1 emphasizes power consumption, while FOM_2 emphasizes resolution. For the first, the smaller the value the “better” the modulator; for the second, the larger the value the “better” the modulator. These figures are not adequate for BPΣΔMs because there the bandwidth is not representative of the speed of the underlying circuitry. Thus, for instance, (2.22) should be replaced by,

$$FOM_1 = \frac{P_w}{2^{ENOB} \times \left(f_n + \frac{B_W}{2} \right)} \cdot 10^{12} \Bigg|_{\text{pJ/(conversion step)}} \quad (2.24)$$

and similar definition replacement applies for (2.23).

2.5 DISCRETE-TIME SIGMA-DELTA MODULATORS

Most of this section is devoted to LPΣΔMs. BPΣΔMs, which are covered with larger details in Chapter 11, are briefly revised by the end of the section.

-
- 18. F. Goodenough, “Analog Techniques of all Varieties Dominate ISSCC”. *Electronic Design*, Vol. 44, pp. 96-111, February 1996.
 - 19. S. Rabii and B. Wooley, “A 1.8-V Digital-Audio Sigma-Delta Modulator in 0.8-μm CMOS”. *IEEE J. of Solid-State Circuits*, Vol. 32, pp. 783-796, June 1997.

2.5.1 Single Quantizer Single-Bit Architectures

Table 2.1 summarizes reported sigma-delta modulator ICs belonging to this family. The simplest among them is the second-order one – see Fig. 2.15 where the DAC block is explicitly shown for convenience. It employs two integrators to feature second-order filtering of the quantization error, thus providing the following in-band error power in the ideal case,

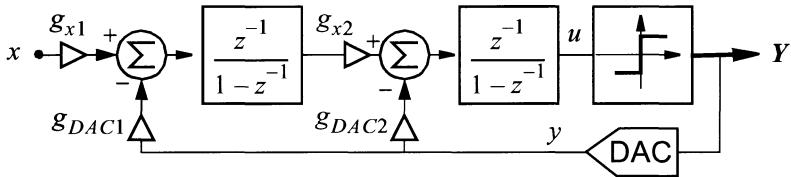
$$P_{\epsilon q} \approx \frac{\Delta^2}{12} \frac{\pi^4}{5M^5} \quad (2.25)$$

Although a simpler structure using just one integrator (the first order sigma-delta modulator) can be built, it is prone to generate repetitive output patterns (which

Table 2.1: Single-quantizer DT $\Sigma\Delta$ modulators, single-bit.

MS means mixed-signal technology, and ST means standard

	DR (bit)	DOR (MS/s)	M	Archit.	Process	Power (mW)	FOM₁	FOM₂ $\times 10^5$
[Send97]	14.33	0.006	128	2nd-order	0.5 μ m MS / 1.5V	0.55	4.46	11.55
[Pelu96]	12	0.0068	74	2nd-order	0.7 μ m ST / 1.5V	0.101	3.63	2.82
[Grilo96]	15.32	0.007	286	2nd-order	0.6 μ m ST / 1.8V	2	6.98	14.62
[Mede97]	16.4	0.0096	256	2nd-order	0.7 μ m ST / 5V	1.71	2.06	104.79
[Bose88]	14.5	0.016	256	2nd-order	3 μ m MS / 5V	12	32.37	1.79
[Tili01]	13	0.016	64	2nd-order	0.25 μ m ST / 1.8V	1	7.63	2.68
[Burm96]	14.16	0.01953125	256	2nd-order	2 μ m MS / 5V	13	36.36	1.26
[Bran91b]	16	0.05	256	2nd-order	1 μ m MS / 5V	13.8	4.21	38.84
[Than97]	13.4	0.1953125	128	2nd-order	1.2 μ m MS / 5V	25.9	12.27	2.20
[Nade94]	13.83	0.002	250	3rd-order	2 μ m MS / 5V	0.94	32.27	1.13
[Au97]	12	0.016	64	3rd-order	1.2 μ m MS / 2V	0.34	5.19	1.97
[Pelu98]	12.5	0.032	48	3rd-order	0.5 μ m ST / 0.9V	0.04	0.22	67.00
[Yama94]	18	0.00002	1600	4th-order	1.2 μ m MS / 5V	1.3	247.96	2.64
[Kert94]	21	0.0008	320	4th-order	3 μ m MS / 10V	25	14.90	351.28
[Kash99]	19.97	0.0008	320	4th-order	0.6 μ m MS / 5V	16	19.47	131.63
[Briga02]	17.1	0.0008	320	4th-order	0.6 μ m MS / 5V	50	444.90	0.79
[Snoe01]	16.65	0.022	64	4th-order	0.5 μ m MS / 2.5V	2.5	1.11	232.29
[Bajd02]	13	0.022	64	4th-order	0.5 μ m MS / 1.8V	1.7	9.43	2.17
[Coba99]	16	0.04	64	4th-order	0.5 μ m ST / 1.5V	1	0.38	428.81
[Maul00]	15.32	0.5	64	5th-order	0.6 μ m ST / 5V	210	10.27	9.94



Weight	[Bose88]	[Yin94]	[Marq98b]	[Mede99a]
g_{x1}, g_{DACP1}	0.5	0.25	1/3	0.25
g_{x2}	0.5	0.5	0.6	1
g_{DACP2}	0.5	0.25	0.4	0.5
Total int. OS / E_r	3.5	2	2.4	2
# unitary caps.	6	11	12	9

Figure 2.15: Second-order Sigma-Delta Modulator.

OS denotes the integrator output swing

2nd-order $\Sigma\Delta$ Ms were proposed in J.C. Candy, "A Use of Double Integration in Sigma-Delta Modulation". *IEEE Transactions on Communications*. Vol. 33, pp. 249-258, March 1985.

makes SNR decrease) and, hence, impractical for achieving performance.

Second-order modulators include four analog coefficients whose selection raises important design issues. Considerations influencing selection of such coefficients include:

- Realizing the targeted noise-shaping function,
 - Precluding the quantizer from overloading,
 - Keeping the state variables (integrator outputs) bounded,
 - Simplifying circuit implementation of such coefficients
- ...

According to Candy (see caption of Fig. 2.15), stability of second-order modulators is guaranteed provided that $g_{DACP2} > 1.25 \cdot g_{x1} \cdot g_{x2}$. The table attached to Fig. 2.15 presents different practical coefficient choices reported in literature, together with the number of unitary capacitors needed for implementation with switched-capacitor circuits, and the maximum swing at the integrator outputs normalized to the reference level, E_r . The latter may become an important issue in low-voltage implementations, where small-swing analog coefficients should be preferred.

The implementation concept underlying Fig. 2.15 can be extended towards L th-order filtering, thus resulting into the modulator structure of Fig. 2.16. The ideal upper value for SNR of this structure is,

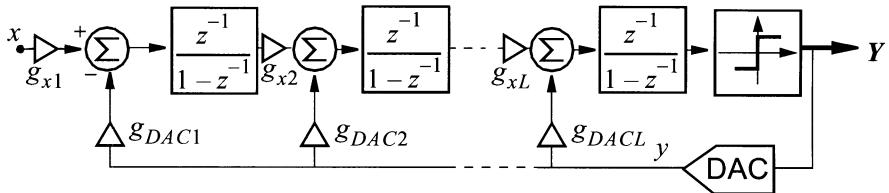


Figure 2.16: Single-loop L th-order Sigma-Delta Modulator.

$$SNR_{\max} \approx [6,02 \cdot L + 3,01] \cdot \log_2(M) + 6,02 \cdot \log_2\left(\frac{\sqrt{2 \cdot L + 1}}{\pi^L}\right) \quad (2.26)$$

However, this upper value cannot be reached in practice due to instability, which, although very involved for exact analysis, can be linked to the large value of the noise transfer function at high frequencies.

High-order $\Sigma\Delta$ Ms are conditionally stable; i.e., they remain stable as long as the input signal and the state variables meet certain requirements. In order to guarantee conditional stability analog coefficients must be properly chosen. Empirical procedures for choosing the coefficient values have been presented elsewhere [OpT93] [Schr93] [Marq98b]. These procedures attempt basically to decrease the high-frequency gain of the noise transfer function, maximizing simultaneously the SNR .

Despite the usage of optimized noise transfer function, instability may appear. In such cases the system can be returned to a stable operating region by one of the following techniques:

- Resetting the integrators to zero or some other initial condition, when unstable operation is detected. The detection of instability can be done at the integrator level, by placing comparators to determine whether an internal state variable has surpassed a certain limit, or by monitoring the length of the series of consecutive pulses at the modulator output. Regarding resetting itself, it may be global (for all the integrators), or local (for some selected integrators) – see for instance [Au97].
- Identifying the maximum amplitudes of the integrator output excursions during stable operation and including limiters to preclude these outputs from taking values larger than these maxima.

In any case, quite significant SNR degradations result.

There are other possibilities for obtaining high-order noise-shaping functions [Ribn91]. The architecture proposed by Lee and Sodini [Lee87] (see Fig. 2.17) is representative of the approach. Assuming a delay in the quantizer, it obtains a

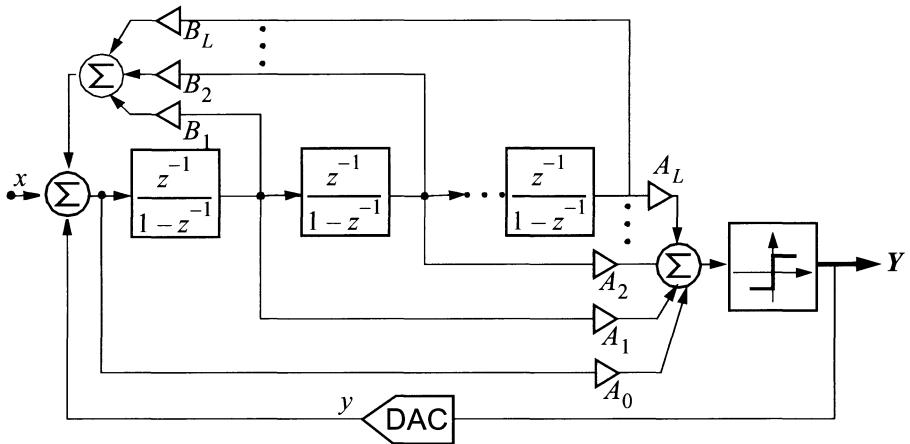


Figure 2.17: Lee-Sodini L th-order $\Sigma\Delta$ modulator

noise transfer function,

$$N_{TF}(z) = \frac{(z-1)^L - \sum_{i=1}^L B_i \cdot (z-1)^{L-i}}{z \cdot \left[(z-1)^L - \sum_{i=1}^L B_i \cdot (z-1)^{L-i} \right] + \sum_{i=0}^L A_i \cdot (z-1)^{L-i}} \quad (2.27)$$

with multiple poles and zeros along the signal-band and much smaller gain out of that band. If all B_i coefficients are zero, $N_{TF}(z)$ has all zeros located at DC. The high-pass function is thus similar to that obtained with a Butterworth or Chebyshev filter. Otherwise, if B_i 's are adjusted to place the zeros in the stop-band, the selectivity of the filter is maximized, and consequently, the in-band quantization noise power minimized, with characteristics similar to an inverse Chebyshev or elliptic filter.

Apart from the Lee-Sodini modulator, also called *interpolative*, there exist other possibilities for obtaining high-order noise-shaping functions, with the characteristics of most used filters [Nors97] [Enge99]. Most of them use feedback or feedforward transmission of the signals in the loop. A drawback of these modulators is the increased complexity of the analog circuitry.

As demonstrated through the results in Table 2.1, modulators covered in this section are not well suited to achieve high-resolution at high-frequencies because they need large oversampling ratios in order to compensate for the *SNR* degradation caused by instabilities.

2.5.2 Single Quantizer Multi-Bit Architectures

Table 2.2 summarizes reported sigma-delta modulator ICs belonging to this family. As to the architectures in previous section, the only difference lies in the usage of a multi-bit quantizer, instead of a single-bit one. This has pros and cons. Among the pros:

- Smaller quantization step, Δ , and hence smaller in-band power of error than 1-bit quantizers
- Weaker nonlinearity than for 1-bit quantizer. Hence, phenomena caused by the nonlinear dynamics, such as idle patterns, tones, dead zones [Enge99], . . . are less notorious.
- Better fitting to the white noise, additive quantization error model than for single-bit quantizers.
- Better stability properties than 1-bit quantizers for given loop filter order.

Generally speaking, the larger the number of bits, the more noticeable these pros, and the closer SNR value can be made to the corresponding ideal value. It is illustrated in Fig. 2.18, taken from [Broo02]. Note that if coefficients are optimized for 1bit, the SNR improves only at 6dB-per-bit rate. This choice is far from optimum. By re-optimizing the coefficients for each new B value much larger improvement is achieved. In such a case, the 2nd-order modulator features, for $B > 3$, larger SNR than the sub-optimum 3rd-order one. By re-optimizing this latter around 30dB SNR enhancement is achieved for $B = 4$, simply due to the improved stability properties. As B increases, the rate of SNR enhancement becomes closer to 6dB-per-bit, suggesting that stability properties improve rap-

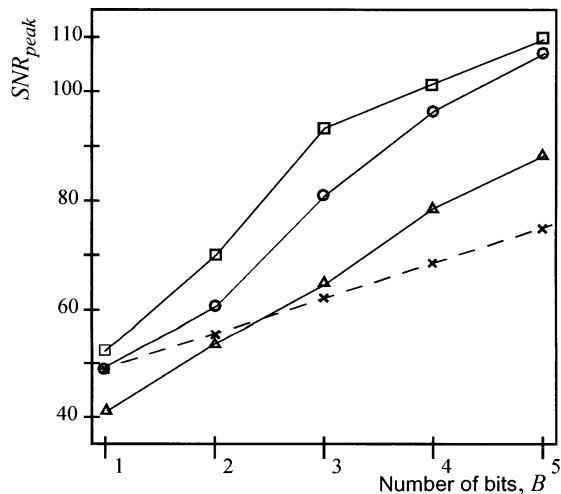


Figure 2.18: SNR_{peak} vs B for different optimized noise transfer functions with $M = 16$.

- 3th-order, optimized for $B=1$
- 3th-order, optimized for each B value
- 2th-order, optimized for each B value
- 5th-order, optimized for each B value

Drawings are taken and depicted from [Broo02]

idly with the number of bits, particularly for higher order modulators.

These considerations suggest that multi-bit quantizers enable reducing the oversampling ratio needed for given resolution, hence being better suited for high-frequency operation. This is actually highlighted by the results presented in Table 2.2. Unfortunately, multi-bit quantization have also drawbacks that overshadow its big advantages; among them:

- Multi-bit quantizers are more difficult to design than single-bit ones.
- Most importantly, they exhibit errors, mainly due to *mismatch*, that have significant impact on the performance of the modulator because some of them are not attenuated by error processing.

Errors associated to multi-bit quantizers are covered in the chapters devoted to Nyquist data converters as well as in Chapter 8. Their impact on $\Sigma\Delta$ Ms is understood with the help of Fig. 2.19. There, new quantizer errors, respectively associated to the A-to-D conversion process and the subsequent D-to-A conversion needed to reconstruct the analog feedback signal, are considered in addition to the quantization error e_q . Note that injections of the ADC error, e_{ADC} , and DAC error, e_{DAC} , happen at different locations within the flow graph. The former is injected into the same path as the quantization error, e_q , so that it is attenuated by the action of feedback. However, the DAC error is injected into the feedback path.

Table 2.2: Single-quantizer DT $\Sigma\Delta$ modulators, multi-bit.

	DR (bit)	DOR (MS/s)	M	Archit.	Process	Power (mW)	FOM₁	FOM₂ x 10⁵
[Nys97]	19	0.0008	512	2nd-ord (3b)	2 μ m MS / 5V	2.175	5.19	252.36
[Chen95]	15.65	0.04	64	2nd-ord (3b)	1.2 μ m MS / 5V	67.5	32.82	3.91
[Sarh93]	15.66	0.041	128	2nd-ord (4b)	2 μ m MS / 5V	--	--	--
[Grilo02]	13	1	32	2nd-ord (4b)	0.35 μ m BiCMOS / 2.7V	11.88	1.45	14.10
[Fogl00]	16.22	0.048	64	2nd-ord (5b)	0.5 μ m ST / 3.3V	68.6	18.72	10.18
[Fogl01]	16.65	0.040	64	2nd-ord (5b)	0.5 μ m ST / 3.3V	70.4	17.11	15.00
[Mille02]	15.32 13.50	0.036 0.4	639 57.5	2nd-ord (6b)	0.18 μ m MS / 2.7V	30 6.47	20.37 5.01 6.47 4.47	
[Mille02]	12.83 11.67	1.25 3.84	18 12	2nd-ord (6b)	0.18 μ m MS / 2.7V	30 50 4.00	3.30 7.42 4.00	5.51 1.38 2.04
[Geer00]	15.8 12.0	2.5 12.5	24	3rd-ord (4b)	0.65 μ m MS / 5V	295 380	2.07 7.42	68.85 1.38
[Bair96]	13.66	0.5	16	4th-ord (4b)	1.2 μ m MS / 5V	58	8.96	3.61
[Kuo02]	13.7 13.0	1.25 2	12	4th-ord (4b)	0.25 μ m ST / 2.5V	100 105	6.01 6.41	5.53 3.19
[Leun97]	19.3	0.096	64	7th-ord (1.5b)	0.8 μ m MS / 5V	760	12.26	131.36

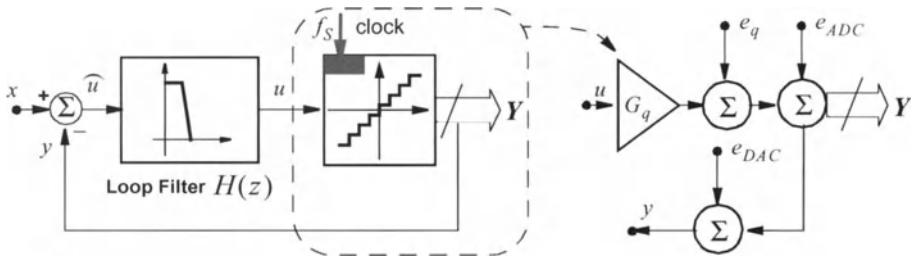


Figure 2.19: Additive error quantizer model.

Thus, this error is processed in the same way as the input itself, which means that it will not be attenuated by the action of feedback. Consequently, the DAC must be designed to reach accuracy levels (particularly regarding *INL* specification) equal to those targeted for the whole system ^{†20}. It is quite challenging due to the impact of mismatch and other circuit imperfections on *INL* values.

Fig. 2.20 shows the architecture typically employed for quantizer design in sigma-delta modulators, where unit elements (transistors, capacitors, resistors, etc.) are employed to reconstruct the analog feedback signal from the digital code. In this structure errors are mainly due to mismatch among the unit elements, according to,

$$\sigma\left(\frac{\Delta y}{y}\right) = \frac{1}{2\sqrt{2}^B} \sigma\left(\frac{\Delta U_e}{U_e}\right) \quad (2.28)$$

where $\sigma(\Delta U_e / U_e)$ is the unit element error. A priori, components can be chosen large enough as to yield the targeted accuracy. However, for resolution levels

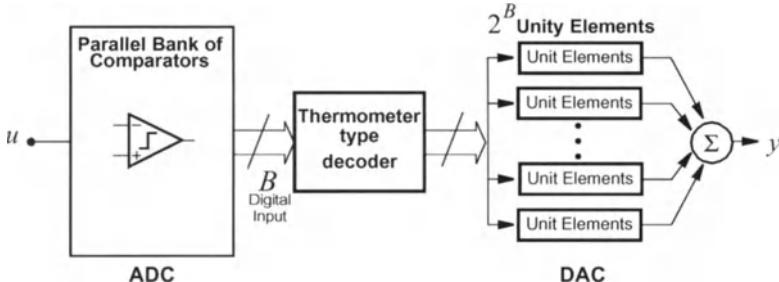


Figure 2.20: Typical quantizer architecture employed for $\Sigma\Delta$ Ms.

-
20. There is another interpretation for the accuracy needed at the DAC. Since low-frequency components of x and y must be practically equal (this is the basis of $\Sigma\Delta$ M operation), errors in y , generated by the DAC, must be smaller than the targeted discrepancies between x and y .

above 12-bits (see Table 2.2) this means prohibitive area occupation. Apart from element trimming, which requires special technologies and which is employed in none of the ICs in Table 2.2, common techniques to overcome the problem of element mismatch include:

- Dynamic Element Matching (DEM), whose basic architecture is depicted in Fig. 2.21. As compared to Fig. 2.20, a block (shuffler) has been added to control the selection of unit elements for a given quantization step. This method relies on the observation that errors in y are systematic because the same unit elements are employed for each quantization step at any time.

Thus, by changing the elements employed to build each quantization level from cycle to cycle, hence in a dynamic way, errors in y can be randomized. This is the role played by the shuffler block.

DEM is implemented in practice by following different methods which may be classified in one of the following categories:

- Purely random selection of the unit elements. Although simple, this has the drawback that the mismatch error becomes unshaped, with the same power inside the band and outside the band.
- Selection of the unit elements following some algorithms in order to shape the mismatch error and reduce its power within the signal band. CLA [Leun92], ILA [Chen95], DWA [Bair95] [Nys97], ... are names for different algorithms reported in literature.
- Noise-shaping DEM, consisting of the incorporation of a filter into the element selection logic in order to achieve high-order shaping of the mismatch noise.
- Correction in the digital domain using look-up tables. Fig. 2.22(a) shows the concept, which basically consists in mapping erroneous digital codes provided by the modulator onto correct ones. This obviously requires calibration for which Fig. 2.22(b) shows a possible implementation. [Sarh93] presents an IC implementation of this concept based on a 2nd-order $\Sigma\Delta M$.

A drawback shared for all multi-bit correction techniques is the large complexity of the circuitry, which in the case of DEM grows as an exponential function of the number of bits B . This has prompted the development of other architectures capable of obtaining high-resolution with low oversampling ratio.

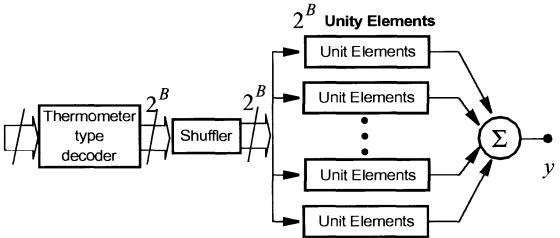


Figure 2.21: Architectural concept for DEM.

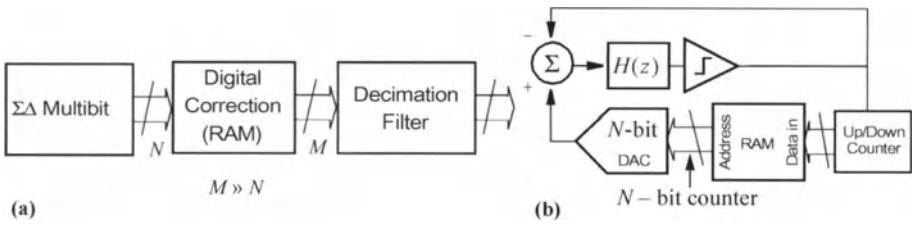


Figure 2.22: Concept for digital calibration.

2.5.3 $\Sigma\Delta$ M Architectures Containing Several Quantizers

All architectures in previous sections contain one quantizer. However, using several quantizers may help to achieve low oversampling ratios without involved analog circuitry, reduced impact of DAC non-linearities and reduced instability problems. Architectures employing several quantizers include cascade $\Sigma\Delta$ Ms (see Chapter 9 of this book), dual quantization $\Sigma\Delta$ Ms [Lesl90], parallel $\Sigma\Delta$ Ms [Galt96] [Wang00], pipeline $\Sigma\Delta$ Ms [Paul99], etc.

2.5.3.1 Dual Quantization

The impact on DAC non-linearity can be attenuated by using dual quantization – illustrated in Fig. 2.23. There, two quantizers are employed: a coarse one, yielding Y_C , and a fine one, yielding Y_F . The digital outputs of these quantizers are substracted to obtain:

$$Y_3 = E_{qF} - E_{qC} \quad (2.29)$$

Then, this signal is digitally processed and the outcome is combined with Y_C to obtain,

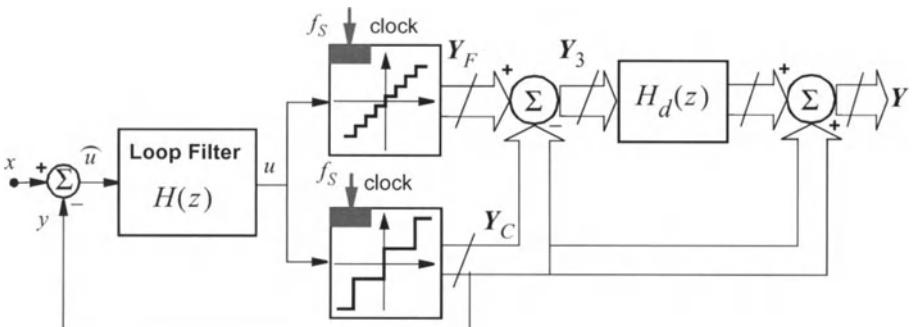


Figure 2.23: Dual quantization concept for error processing in open loop.

T.C. Leslie and B. Singh, “An Improved Sigma-Delta Modular Architecture”. *Proc. of IEEE Int. Symp. on Circuits and Systems*, pp. 372-375, May 1990.

$$Y = S_{TF}(z) \cdot X(z) + N_{TF}(z) \cdot E_{qC}(z) + (E_{qF} - E_{qC}) \cdot H_d(z) \quad (2.30)$$

This equation shows that the coarse quantization error can be nullified by choosing $H_d(z) = N_{TF}(z)$. Thus, only the fine quantization error remains. Note also from Fig. 2.23 that the DAC-reconstructed feedback signal corresponds to coarse quantization. Hence, the impact of DAC non-linearity is smaller than in the conventional architecture which employs only one quantizer.

2.5.3.2 Cascade DT $\Sigma\Delta$ Modulators

Fig. 2.24 illustrates the concept of cascade $\Sigma\Delta$ Ms. It shows two modulators connected in a cascade, such that the quantization error, e_{q1} , of the first stage is processed by the second one. Since this quantization error appears at both Y_1 (as such) and Y_2 (after being processed by the second stage) it can be annulled in the digital domain by properly processing and combining these two signals. After such processing, realized by blocks $H_{d1}(z)$ and $H_{d2}(z)$ in Fig. 2.24, only the quantization error of the second quantizer remains, and it is shaped by the combined transfer function,

$$N_{TF}(z) = N_{TF1}(z) \cdot N_{TF2}(z) \quad (2.31)$$

whose order is obviously larger than those of the component stages; for instance, 3rd-order is obtained by cascading a 2nd-order stage and a 1st-order one. Obviously, expressions of $H_{d1}(z)$ and $H_{d2}(z)$ must be properly chosen to annul the first stage quantization error.

Cascade modulators have similarities with dual quantization architectures in the sense that they all rely on processing in digital domain several replicas of one quantization error signal to nullify its overall contribution. The concept of cascade

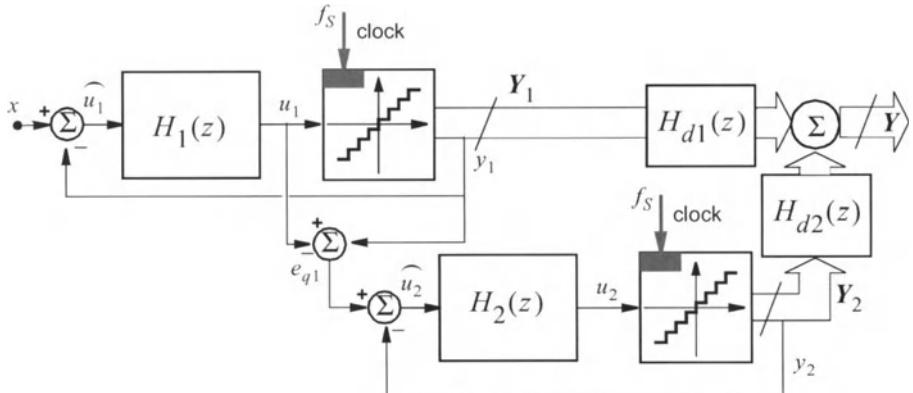


Figure 2.24: Concept of cascade $\Sigma\Delta$ M.

modulators is extensible to whatever number of stages ^{†21}, and applies as well if the second stage is either a simple quantizer or some types of Nyquist converter [Broo97] [Romb01].

Fig. 2.25 shows a generic cascade $\Sigma\Delta M$ architecture where all stages are assumed to contain multi-bit quantizers. Note that there is no inter-stage feedback. Feedback loops are local; i.e., internal to the stages. Hence, provided that only 1st- and/or 2nd-order stages are used, stability properties are simple and do not result into significant *SNR* degradation. Moreover, Fig. 2.25 may implement a high-order noise-shaping function.

Actually, analysis of Fig. 2.25 assuming perfect cancellation, shows that only the signal $X(z)$ and the last-stage quantization error $E_N(z)$ remain at the z -domain modulator output, yielding:

$$Y(z) = z^{-L} \cdot X(z) + d \cdot (1 - z^{-1})^L \cdot E_{qN}(z) \quad (2.32)$$

where,

$$L = L_1 + L_2 + \dots + L_N \quad (2.33)$$

and L_k , $k = 1 \dots N$, is the order of the k -th stage in the cascade. Hence, the filtering realized on the last stage corresponds to that of L th-order $\Sigma\Delta M$ except for a scalar d , equal to the inverse of the product of the integrator weights in the cascade. In order not to prematurely overload the modulator, signal must be properly scaled down as transmitted from one stage to the next one. This results in a value of d larger than unity, which means an amplification of the last-stage quantization

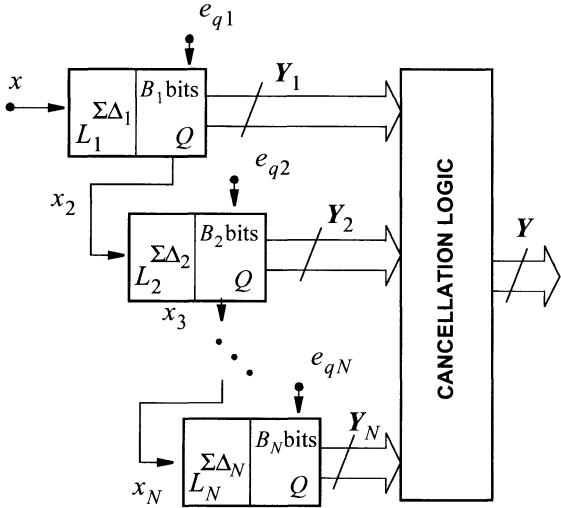


Figure 2.25: Conceptual cascade $\Sigma\Delta M$.

21. There is a practical limitation on the number of stages posed by circuit non-idealities such as mismatch and other circuit errors [Rio02a]. These circuit errors produce incomplete cancellation and hence leakage of the nominally null quantization errors. See Chapter 9.

error, thus degrading the effective resolution. However, this loss can be reduced to only one bit ($d = 2$), no matter which the modulator order is.

Cascade $\Sigma\Delta$ Ms are very well suited to achieve high-resolution with low oversampling ratios [Rio02a]. In practice, they have been implemented either using only single-bit quantizers (see Table 2.3) or combining single- and multi-bit quantizers (see Table 2.4). Specially appealing is the case where multi-bit quantization is employed only in the last stage since this relaxes the multi-bit DAC linearity requirement. In such a case, the non-linearity error is not injected into the modulator input but high-pass filtered, so that most part of the DAC error power is pushed out of the signal band. Analysis shows,

$$\begin{aligned} Y(z) &= z^{-L} \cdot X(z) + d \cdot (1 - z^{-1})^L \cdot E_{qN}(z) + \\ &= + d \cdot (1 - z^{-1})^{(L - L_N)} \cdot E_{DAC}|_N(z) \end{aligned} \quad (2.34)$$

where $E_{DAC}|_N$ captures the integral non-linearity error produced in the last-stage DAC. Note that this error is shaped by a function of order $L - L_N$ (i.e., the overall modulator order minus the order of the last stage). Hence, correction/calibration can be avoided to reduce the impact of DAC non-linearity errors [Rio02a].

Consider for illustration purposes the 2-2-2 cascade in Fig. 2.26; digital trans-

Table 2.3: Cascade DT $\Sigma\Delta$ modulators, single-bit

	DR (bit)	DOR (MS/s)	M	Archit.	Process	Power (mW)	FOM₁	FOM₂ x 10⁵
[Gome00]	16.65	0.044	128	2-1	0.6 μ m MS / 3V	22	4.86	52.88
[Will94]	17	0.05	128	2-1	1 μ m MS / 5V	47	7.17	45.62
[Yin93]	15.7	0.32	64	2-1	1.2 μ m ST / 5V	65	3.82	34.82
[Rabi97]	16.1	0.05	80	2-1	0.8 μ m MS / 1.8V	2.5	0.71	246.29
[Wang01]	18.1	0.025	64	2-2	0.6 μ m MS / 5V	75	10.68	65.68
[Rito94]	16.15	0.044	64	2-2	1.2 μ m MS BiCMOS / 5V	102	31.82	5.72
[Fuji97]	18.15	0.048	128	2-2	0.7 μ m MS / 5V	500	35.91	20.17
[Miao98]	14.82	0.05	64	2-2	3 μ m MS / 5V	74	51.03	1.42
[Olia02]	13.5	0.36	36	2-2	0.4 μ m MS / 1.8V	5	1.20	24.12
[Rebe90]	15	0.18	64	1-1-1	1.5 μ m MS / 5V	76	12.89	6.35
[Yin94]	15.82	1.5	64	2-1-1	2 μ m MS BiCMOS / 5V	180	2.07	69.61
[Marq98a]	14.8	2.0	24	2-1-1	1 μ m MS / 5V	230	4.03	17.66
[Geer99]	15	2.2	24	2-1-1	0.5 μ m MS / 3.3V	200	2.77	29.48
[Mori00]	14	2.2	24	2-2-2	0.35 μ m MS / 3.3V	150	4.16	9.83
[Yoon98]	15.3	0.064	16	2-1-1-2	2 μ m MS / 6.6V	79	30.60	3.29

Table 2.4: Cascade DT $\Sigma\Delta$ modulators, multi-bit.

	DR (bit)	DOR (MS/s)	M	Architecture	Process	Power (mW)	FOM₁	FOM₂ $\times 10^5$
[Broo97]	14.5	2.5	8	2-0(5b)	0.6 μ m MS / 5V	550	9.49	6.09
[Bran91a]	12	2.1	24	2-1(3b)	1 μ m ST / 5V	41	4.77	2.14
[Lamp01]	13	1.5625	32	2-2(3b)	0.35 μ m MS/2.5V	50	3.91	5.23
[Mori00]	13	2.2	24	2-2(5b)	0.35 μ m MS/3.3V	99	5.49	3.72
[Gupta02]	14.6	2.2	29	2-1-1(2b)	0.35 μ m ST/3.3V	180	3.29	18.81
[Mede99b]	13	2.2	16	2-1-1(3b)	0.7 μ m ST / 5V	55	3.05	6.70
[Rio02b]	13.7 13.0	2.2 4.4	32 16	2-1-1(3b)	0.25 μ m ST/2.5V	71.7	2.45 1.99	13.56 10.28
[Rio01]	13	2.2	16	2-1-1(4b)	0.35 μ m ST/3.3V	78.2	4.34	4.71
[Fuji00]	15	2.5	8	2(4b)-1(4b)-1(4b)	0.5 μ m MS / 5V	105	1.28	63.81
[Vleu01]	15	4	16	2(5b)-2(3b)-1(3b)	0.5 μ m MS/2.5V	150	1.14	71.47
[Feld98]	13	1.4	16	2-2-2(1.5b)	0.7 μ m MS/3.3V	81	7.06	2.90
[Dedi94]	14.25	0.2	16	2(1.5b)-2(1.5b)-2(1.5b)	1.2 μ m MS / 5V	40	10.26	4.74

fer functions and expressions for the analog and digital coefficients associated to this architecture are included in the table attached to the figure. It yields,

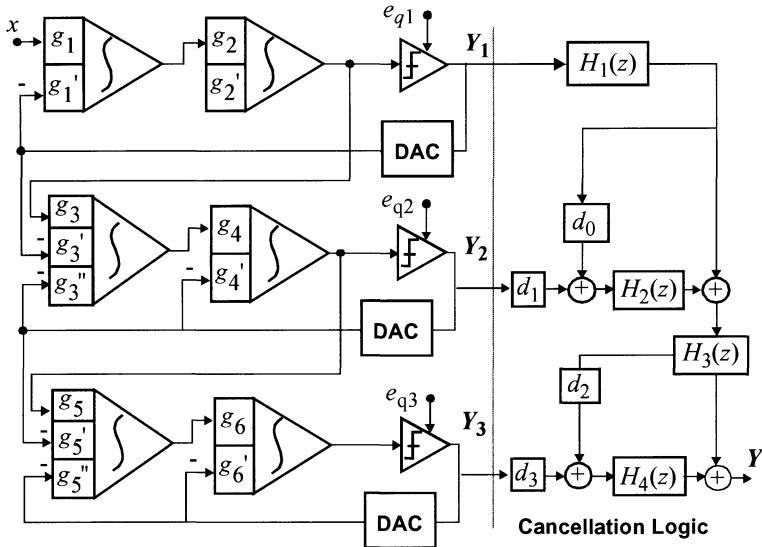
$$P_{\epsilon_q} \approx d^2 \cdot \frac{\Delta^2}{12} \cdot \frac{\pi^{12}}{13M^{13}} + d^2 \cdot \sigma_{DAC}^2 \cdot \frac{\pi^7}{9M^9} \quad (2.35)$$

where σ_{DAC}^2 represents the power of the DAC error e_{DAC} , which can be estimated as $\sigma_{DAC}^2 \approx (\Delta^2/2) \cdot (INL/100)^2$, with INL being the DAC integral non-linearity expressed in %FS.

Similar to any other high-order modulator, proper selection of the analog coefficients is crucial for minimum SNR degradation and simplest circuit implementation. Table 2.5 shows alternative coefficient sets for this structure. Those in column **A** are chosen to minimize the systematic loss of resolution caused by the digital coefficient d in (2.34), while those in columns **B** and **C** contain other sets

Table 2.5: Alternative Coeffs. for the 6th-order 2-2-2 cascade architecture.

A	B	C
0.25, 0.25, 0.5, 0.25	0.2, 0.2, 0.5, 0.25	0.5, 0.5, 0.5, 0.5
0.5, 0.25, 0.25	-0.5, -0.1, 0.2	0.25, 0, 0.5
0.5, 0.25, -	0.5, 0.25, -	0.5, 0.5
0.5, 0.0625, 0.25	-1, -0.2, 0.2	0.25, 0, 0.5
0.5, 0.25	0.4, 0.2	0.5, 0.5



Digital	Digital/Analog	Analog
$H_1(z) = z^{-2}$	$d_0 = \frac{g_3}{g_1 g_2 g_3} - 1$	$g_1' = g_1$
$H_2(z) = (1-z^{-1})^2$	$d_1 = \frac{g_3''}{g_1 g_2 g_3}$	$g_2' = 2g_1' g_2$
$H_3(z) = z^{-2}$	$d_2 = 0$	$g_4' = 2g_3'' g_4$
$H_4(z) = (1-z^{-1})^4$	$d_3 = \frac{g_5''}{g_1 g_2 g_3 g_4 g_5}$	$g_5' = g_3'' g_4 g_5$
		$g_6' = 2g_5'' g_6$

Figure 2.26: 6th-order 3-stage $\Sigma\Delta M$.

proposed for the 2-2-2 cascade [Feld98] [Mori00]. Depending upon the chosen coefficients, the *SNR* is enhanced by more than 12dBs for an oversampling ratio of $M = 16$.

As already mentioned, the concept of cascade applies as well if the second stage is either a quantizer or some type of Nyquist converter. For instance, Fig. 2.27 shows a $\Sigma\Delta$ -pipeline converter architecture reported in [Broo97] and validated through a 14.5bits a 2.5MS/s IC prototype – see Table 2.4. A drawback peculiar to this and, in general, to all cascade architectures is due to incorrect cancellation of early stages quantization errors which can be confronted by using online correction strategies [Kiss00].

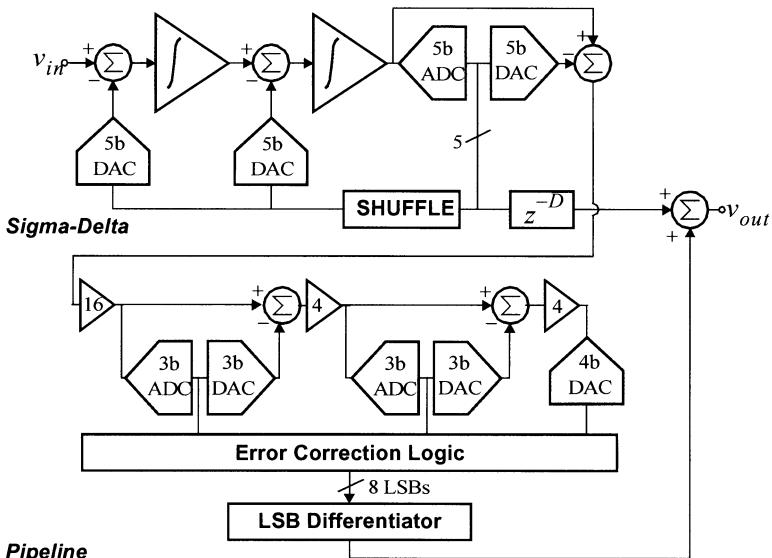


Figure 2.27: $\Sigma\Delta$ -pipeline modulator concept according to [Broo97].

2.5.3.3 Parallel Sigma-Delta Architectures

Fig. 2.28 shows a conceptual block diagram for parallel $\Sigma\Delta$ modulators. It is composed of K modulators and a bank of digital filters and multiplexors. Assuming that modulator components operate at a sampling frequency f_S , the sampling frequency of the composed architecture goes up to $K \cdot f_S$ [Aziz93] [Corn94] [Esh96] [Galt95] [Galt96] [Khoi97] [King98] [Koza00] [Wang00].

There are three basic approaches for multiplexing the signals coming from the modulators [Esh96]:

- Frequency Domain Multiplexing, where the signal band is split into small sub-bands which are digitized using BP $\Sigma\Delta$ s [Aziz93] [Corn94].
- Time Domain Multiplexing, where the input signal is sampled at high-fre-

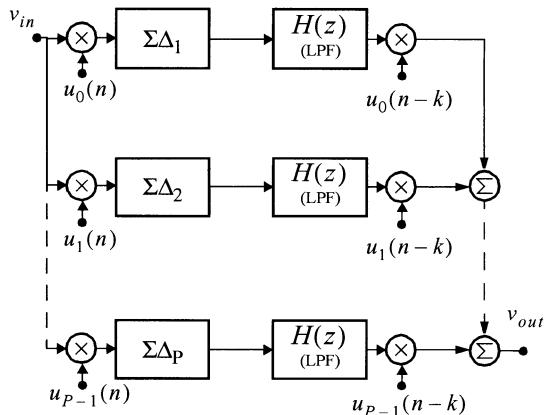


Figure 2.28: Conceptual Block Diagram for a Parallel $\Sigma\Delta$ Modulator.

quency, sub-samples are distributed and modulated at low rate by each of the parallel paths and digital outputs are combined by following the principles of multi-rate digital systems [Khoi97] [Koza00] [Wang00].

- Code Division Multiplexing, where the input signal is multiplied by a different Hadamard sequence per channel ($u_k(n)$ in Fig. 2.28). Modulator outputs are then lowpass filtered, multiplied by delayed versions of the corresponding Hadamard sequences ($u_k(n - k)$ in Fig. 2.28), and finally combined in digital domain [Galt95] [Galt96] [King98].

These strategies are not yet mature for IC implementation; very few IC prototypes have been reported [King98].

2.5.4 BP DT $\Sigma\Delta$ modulators

BP $\Sigma\Delta$ Ms differ from LP ones in that the loop filter is bandpass type instead of lowpass type (see Fig. 2.13). It is hence characterized by two parameters, namely the bandwidth B_W and the central frequency f_n . This latter can be located at different points in the interval comprised between DC and $f_S/2$. Selecting its actual location defines a trade-off between anti-aliasing filtering and image-rejection filtering. A common choice is using $f_n = f_S/4$ (see Chapter 11). Linked to this choice, BP $\Sigma\Delta$ M architectures are obtained from LP $\Sigma\Delta$ M ones by applying the DT lowpass to bandpass transformation, namely by making,

$$z^{-1} \rightarrow -z^{-2} \quad (2.36)$$

In practice, it means replacing the integrators employed in LP $\Sigma\Delta$ Ms by resonators and, consequently, the zeroes of the lowpass noise transfer function are moved from DC to $f_S/4$. It is illustrated in Fig. 2.29 for a second-order modulator. This type of transformation is underlying in most of BP $\Sigma\Delta$ M ICs reported in literature

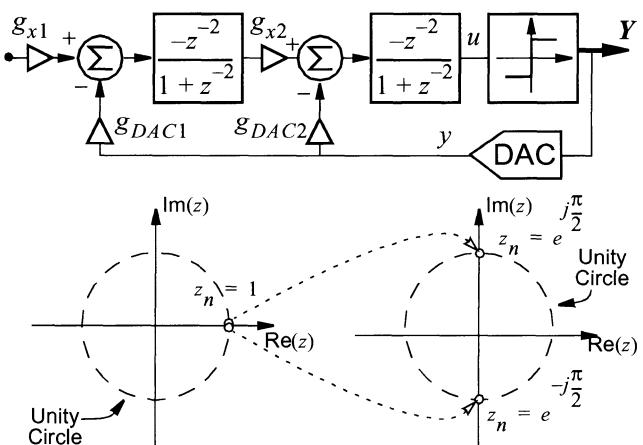


Figure 2.29: Illustrating LP-to-BP transformation for a second-order modulator.

because it results into modulators which keep all features of the original lowpass structures. Table 2.6 summarizes main features of BPΣΔM ICs, based on the transformation in (2.36), reported in literature.

Similar to lowpass modulators, it may be convenient for bandpass to directly address the synthesis of the noise transfer function through proper placement of its poles and zeroes. For illustration purposes, Fig. 2.30(a) shows a fourth-order BPΣΔM, composed of a cascade of resonators, designed using this method [Jant93]. Fig. 30(b) shows the pole/zero location of $S_{TF}(z)$ and $N_{TF}(z)$. Observe that the zeroes of $N_{TF}(z)$ are placed around $f_S/4$. In this case $S_{TF}(z)$ realizes a bandpass function which acts as an anti-aliasing filter – see Fig. 2.30(c). Another

Table 2.6: BPΣΔM ICs based on $z^{-1} \rightarrow -z^{-2}$ transformation.

	DR (bit)	f_S (MHz)	f_n B_w (MHz)	Archit. of the LP prototype	Process	Power (mW)	FOM₁	FOM₂ $\times 10^6$
[Baza98]	6.7	40	20 1.25	2nd-ord	0.5μm MS / 5V	65	30.31	0.09
[Cheu01]	6.7	42.8	10.7 0.2	2nd-ord	0.35μm ST / 1V	12	10.69	0.24
[Rosa00]	11.5	6.52	1.63 0.01	4th-ord	0.8 μm ST/ 5V	60	12.67	5.71
[Long93]	15	7.2	1.8 0.03	4th-ord	1μm MS / 5V	--	--	--
[Song95]	9	8	2 0.03	4th-ord	2μm MS / 3.3V	0.8	0.78	16.48
[Park99]	12.2	20	5 0.2	4th-ord	0.65μm ST / 4V	180	7.50	15.66
[Corm97]	9.5	1.25	0.25-.375 0.00625	4th-ord	2μm MS / 5V	--	--	--
[Salo02]	11.7	80	20 0.27	4th-ord	0.35μm -DP / 3V	56	0.84	99.34
	6.7	80	20 3.84				24.57	0.11
[Taba99]	13	80	20 1.25	6th-ord	0.25μm MS / 2.5V	90	0.53	383.86
[Andr96]	8	8	2 0.064	6th-ord	0.5μm DP / 3.3V	8	15.38	0.42
[Haira96]	11.7	13	3.25 0.2	4-2	0.8μm DP / 3V	14.4	1.29	64.27
[Baza99]	9.4	68	17 1.25	4-4	0.6μm DP / 3V	48	4.03	4.18
[Ueno02]	12.6	10	0.566 0.25	2-2(3b)	0.25μm DP/ 2.5V	77	17.95	8.63

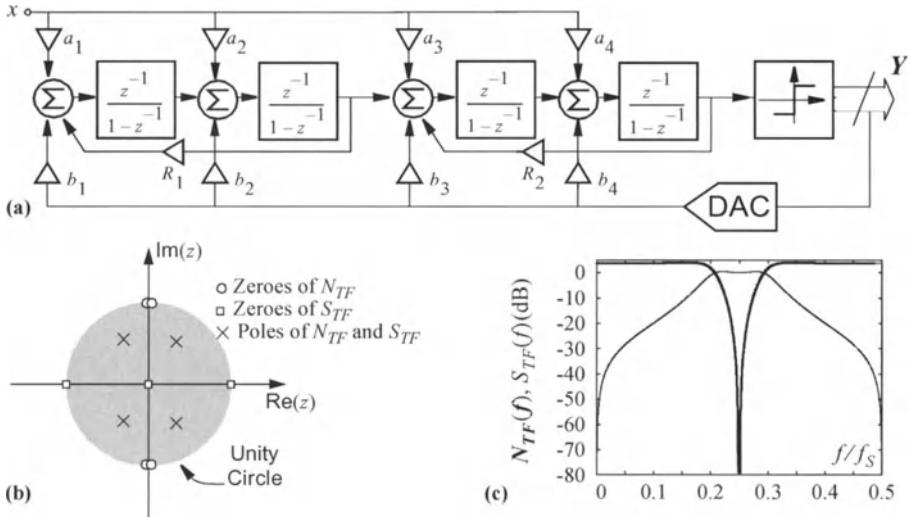


Figure 2.30: An example of a fourth-order BPΣΔM obtained using the direct $N_{TF}(z)$ method [Jant93]. (a) Modulator architecture. (b) Pole/zero location of $S_{TF}(z)$ and $N_{TF}(z)$. (c) $S_{TF}(f)$ and $N_{TF}(f)$ vs. frequency.

Table 2.7: Other BPΣΔM CMOS ICs.

	DR (bit)	f_s (MHz)	f_n B_w (MHz)	Architecture	Process	Power (mW)	FOM₁	FOM₂ x 10⁶
[Jant93]	10.2	1.825	0.455 0.008	4th, Opt. N_{TF}	3μm DP / 5V	210	388.96	0.08
[Liu97]	11.8	0.827	0.413 0.002	4th, Opt. N_{TF}	2μm DP / 5V	--	--	--
[Andr98]	9.2	4	1 0.2	3th(3b), Opt. N_{TF}	0.5μm DP / 3V	19	29.37	0.50
[Chua98]	13	0.5	0.125 0.0005	6th, Opt. N_{TF}	2μm DP / 5V	--	--	--
[Toni99]	12.7	42.8	10.7 0.2	6th, Opt. N_{TF}	0.35μm SP / 3.3V	80	1.11	149.19
[Cusi00]	12	37.05	10.7 0.2	6th, Opt. N_{TF}	0.35μm SP / 3.3V	116	2.62	38.99
[Jant97]	10.8	10	3.75 0.2	4th, Quadrature	0.8μm SP / 5V	130	18.94	2.35
[Ong97]	12.2	80	20 0.2	4th-ord	0.6μm ST / 3.3V	72	0.76	154.26
[Taba00]	12	64	16 2	6th, 2-path	0.25μm SP / 2.5V	110	1.58	64.72

approach for the realization of BPΣΔMs uses the *N-path* design technique [Ong97] [Taba00]. Finally, the concept of BPΣΔMs can be extended into that of quadrature BPΣΔM, which embeds complex bandpass filters in the loop [Jant97]. Table 2.7 shows the main features of several BPΣΔM ICs employing strategies others than the transformation in (2.36).

2.6 CONTINUOUS-TIME SIGMA-DELTA MODULATORS

All previous ΣΔM architectures employ discrete-time loop filters and hence require that the sampling operation is realized at the modulator input – see Fig. 2.31(a). However, this is not mandatory. Instead, the sampling operation can be realized before quantization, and the loop filter can be *continuous-time* – see Fig. 2.31(b). Thus, the output signal is discrete-time, the input signal is continuous-time and a discrete-to-continuous time transformation is needed to create the feedback signal $y(t)$. The process of reconstructing this signal may have significant impact on the overall modulator performance^{†22}. Actually, a proper choice of the shape of the DAC output pulses plays a significant role on

the overall performance of CT ΣΔMs. Three basic choices are depicted in Fig. 2.31(c); respectively called, from left to right, Non-Return-to-Zero (NRZ), Return-To-Zero (RTZ) and Half-delay return-to-Zero (HZ).

Although they combine DT and CT signals, modulators fitting into the archi-

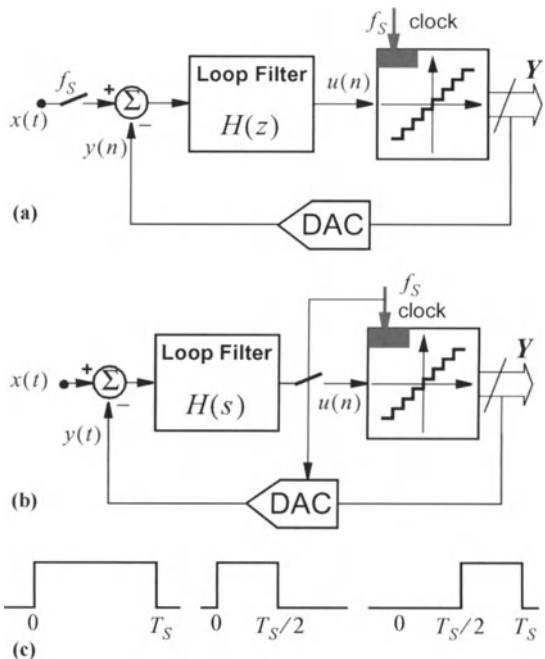


Figure 2.31: Illustrating architectural differences between: (a) DT; (b) CT ΣΔMs; (c) Some basic DAC output shapes.

22. Bear in mind that in-band components of x and y must be practically equal. Hence, the power of error caused by the reconstruction process must be kept smaller than the power of error corresponding to the resolution level specified for the whole modulator.

ture of Fig. 2.31(a) are generically called Continuous-Time (CT) $\Sigma\Delta$ modulators. CT $\Sigma\Delta$ s are more difficult to analyze than pure DT modulators due to the combination of both types of dynamics with hard non-linearities. Theoretical and practical issues associated to CT $\Sigma\Delta$ s have been addressed in several recent monographs [Bree01] [Cher00] [Enge99] [Shoa95]. Their claimed pros include:

- Extra anti-aliasing filter can be avoided. Instead, the CT loop filter $H(s)$ can be exploited to this purpose.
- Less impact of errors caused during sampling operation. This is a consequence of the fact that sampling is not realized at the input, where its error cannot be attenuated, but within the modulator loop.
- No “settling” error at the loop filter circuitry. In DT circuits, signals must settle to their steady-state values according to the dynamic features of the circuitry [Rio02a]. Complete settling would require infinite time. In practice, settling process must last until the influence of the circuit time-constants becomes negligible.
- Larger operation speed. This is inherent in the operation of CT circuits, where circuit dynamics is not a parasitic (like in DT circuits), but a design primitive.
- Absence of $k \cdot T/C$ noise at the input capacitors – a by-product of the location of the sampling operation.
- Reduced digital noise coupling.
- etc.

While their cons include:

- Very involved dynamic due to the combination of non-linearity, continuous-time and discrete-time.
- Larger impact of circuit non-linearities. This is inherent in CT circuits. These circuits require voltage-to-current transformation, and this yields non-linear errors. These errors may have significant influence, particularly those appearing at the very input because they are not attenuated by the modulator feedback loop.
- Time constant tuning is needed for correct loop filtering. This is also an inbuilt feature of CT circuits, where time constants are given as the product of two statistically uncorrelated parameters.
- Time uncertainty (“jitter”) also plays a significant role because it maps directly onto errors in the feedback signal $y(t)$.
- Large sensitivity to loop delay, for the same reason.
- etc.

Usually the design of CT $\Sigma\Delta M$ is done in discrete time domain. To this purpose, an equivalent discrete-time loop filter must first be found. Consider Fig. 2.32(a). It can be redrawn as shown in Fig. 2.32(b), where the loop filter action is separated for the input and the signal paths, and the sampling operation is placed in front of the input summing point. In Fig. 2.32(b), a DT signal path is identified from the quantizer output to the quantizer input. The transfer function of this signal path plays similar role than $H(z)$ in Fig. 2.31(a). As the signal travels across this path, it is first transformed into a CT signal by the DAC. Assuming the DAC shapes in Fig. 2.31(c), the following feedback waveform is obtained in each period,

$$y(t) = y(n) \cdot p_{DAC}(t) = y(n) \cdot [\delta_1(t - p_1 \cdot n \cdot T_S) - \delta_1(t - p_2 \cdot n \cdot T_S)] \quad (2.37)$$

where $\delta_1(\bullet)$ denotes the unit step waveform, $y(n)$ is the quantized version of $u(n)$, and p_1 and p_2 are given in footnote 23^{†23} for the three basic DAC shapes. This signal $y(t)$ is processed by the continuous-time filter to obtain $z(t)$, so that the overall transfer function (DAC shaping + CT filtering) is expressed as:

$$Z(s) = H(s) \cdot \frac{1}{s} \left(e^{-s \cdot p_1 \cdot T_S} - e^{-s \cdot p_2 \cdot T_S} \right) \quad (2.38)$$

Finally, the signal $z(t)$ is sampled, which, in time domain, can be expressed as:

$$z(n) = z(t) \cdot \sum_{n=0, \infty} \delta_0(t - n \cdot T_S) \quad (2.39)$$

23. This transfer function is valid for the three shapes in Fig. 2.31(c) by making:

$$p_1 = 0 \quad p_2 = 1 \quad , \text{ for NRZ}$$

$$p_1 = 0 \quad p_2 = 1/2 \quad , \text{ for RZ}$$

$$p_1 = 1/2 \quad p_2 = 1 \quad , \text{ for HRZ}$$

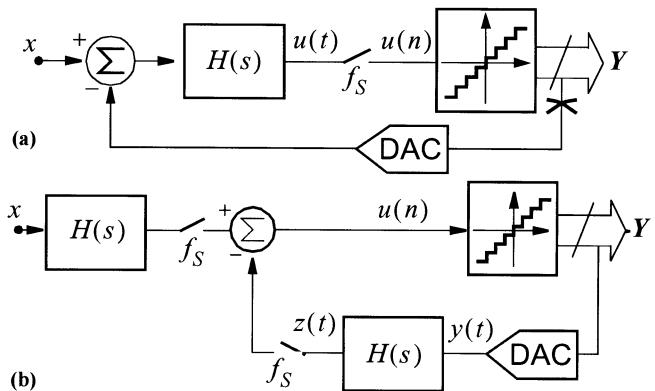


Figure 2.32: Redrawing a CT $\Sigma\Delta M$ for CT-DT transformation.

where $\delta_0(\bullet)$ is the Dirac delta. By combining previous expressions, the transfer function of the overall feedback path is calculated as:

$$H(z) = Z\left\{L^{-1}[DAC(s) \cdot H(s)] \cdot \left[\sum_{n=0}^{\infty} \delta_0(t - n \cdot T_S) \right]\right\} \quad (2.40)$$

where $Z(\bullet)$ stands for the z -transform operator, $L(\bullet)$ stands for the Laplace-transform operator.

By way of example, Table 2.8 shows DT transfer functions corresponding to a CT resonator transfer function for the three basic DAC shapes included in Fig. 2.31(c). Based on the equivalent DT loop filter transfer function, the procedure to design CT $\Sigma\Delta$ Ms consists of, first, matching this equivalent filter with a reference DT loop filter chosen to fulfill specs; then, solving for the coefficients of the CT filter; and finally implementing this filter by circuits. Careful choice of the CT filter structure is needed to have sufficient degrees of freedom to implement the reference DT loop filter [Enge99].

An alternative method for the design of the loop filter uses the desired noise transfer function as a starting point, just in the same manner as for the discrete time case. An Inverse Chevyshev distribution of the noise transfer function zeroes has advantages in terms of SNR and stability. Once the desired noise transfer function has been chosen, the necessary loop filter can be derived from the linearized model [Bree01]. There are two major drawbacks to this method:

- Previous knowledge of discrete time modulators is not reused. This rises questions about stability.
- Simulations are harder due to the fact that every simulation has to be done in continuous time.

These, and similar procedures have been applied for the implementation of a large variety of CT $\Sigma\Delta$ M ICs. Table 2.9 and Table 2.10 summarize main features

Table 2.8: Illustrating CT-to-DT transform.

DAC	$H(z)$	$H(s)$
NRZ	$\frac{z^{-1} \cdot (1 - z^{-1})}{1 + z^{-2}}$	
RZ	$\frac{\left(1 - \frac{\sqrt{2}}{2}\right) \cdot z^{-1} - \left(\frac{\sqrt{2}}{2} \cdot z^{-2}\right)}{1 + z^{-2}}$	$\frac{\omega_o \cdot s}{s^2 + \omega_o^2}$
HRZ	$\frac{\frac{\sqrt{2}}{2} \cdot z^{-1} - \left(\left(1 - \frac{\sqrt{2}}{2}\right) \cdot z^{-2}\right)}{1 + z^{-2}}$	

of CT $\Sigma\Delta M$ ICs reported in literature – only CMOS technologies.

Table 2.9: Continuous-Time LP $\Sigma\Delta M$ CMOS ICs.

	SNR DR (bit)	f_S (MHz)	B_w (MHz)	Architecture	Process	Power (mW)	FOM₁	FOM₂ × 10⁵
[Luh98a]	8	50	1	2nd-ord	2μm CMOS / 5V	15	29.30	0.02
[Luh98b]	9.6	50	1	2nd-ord	2μm CMOS / 5V	16.6	10.70	0.18
[Lin99]	10.5	80	2.5	2nd-ord	1.2μm CMOS / 3V	12	1.66	2.18
[Hall92]	10	150	0.585	2nd-ord	2μm CMOS / 5V	--	--	--
[Zwan99]	10.4	10	0.5	2nd-ord	0.5μm CMOS / 5V	7.2	5.33	0.63
[Gerk02]	11.33	2.4	0.025	3rd-ord	0.5μm CMOS / 1.5V	0.25	1.94	3.31
[Hall93]	10	150	—	4th-ord	2μm CMOS / 5V	--	--	--
[Zwan96]	13	0.512	0.004	4th-ord	0.5μm CMOS / 2.2V	0.2	3.05	6.70
[Veld02]	11.3	153.6	2	4th-ord	0.18μm CMOS/1.8V	6.6	0.65	9.62
[Redm94]	13	2.8	0.04375	4th-ord	1.6μm CMOS / 5V	--	--	--
[Luh00]	10	400	3.1	5th-ord	0.6μm CMOS / 3.3V	16	2.52	1.01

Table 2.10: Continuous-Time BP $\Sigma\Delta M$ CMOS ICs.

	SNR DR (bit)	f_S (MHz)	f_n B_w (MHz)	Architecture	Process	Power (mW)	FOM₁	FOM₂ × 10⁵
[Hsu00]	6.7	280	70 0.2	2nd-ord	0.5μm CMOS 2.5V	39	5.35	0.048
[Bree00] ^a	13.3	13	13 0.1	4th-ord	0.35μm CMOS 2.5V	1.8	0.89	2.82
[Tao99] ^b	7.2	400	100 0.2	4th-ord	0.35μm CMOS 3.3V	165	11.21	0.033
[Enge99]	10.8	40	10.7 0.2	6th-ord	0.5μm CMOS 5V	60	3.12	1.42
[Zwan00]	13.3	21.07	10.7 0.2	5th-ord	0.25μm CMOS 2.5V	11	0.10	249.27

a. In this case, the design includes a mixer and the system is BP. However, considering only the converter, it is better classified as a LP.

b. The actual system is a quadrature converter. In order to calculate FOM1 and FOM2, the total power has been divided by two.

2.7 ABOUT THE INFLUENCE OF CIRCUIT NON-IDEALITIES

Assume ideal circuit components are available to build $\Sigma\Delta$ Ms. Then, resolution is limited only by the quantizer granularity, the oversampling ratio and the filtering of the quantization error. However, in practice other limitations arise due to the unavoidable circuit component non-idealities. They manifest as extra in-band error power terms. While the ideal error power contains only the nominal term $P_{\epsilon q}$ (see for instance (2.14)), the actual error power includes many other terms,

$$P_{\epsilon T} = P_{\epsilon q} + \Delta P_{\epsilon q} + P_{\epsilon DAC} + P_{\epsilon ADC} + P_{\epsilon th} + P_{\epsilon Settling} + P_{\epsilon Jitter} + \dots \quad (2.41)$$

These other terms can be classified into three major groups, namely:

- Quantization error power enlargement, represented by $\Delta P_{\epsilon q}$, due to modifications of the noise transfer function caused by circuit non-idealities.
- Additional error power terms generated by non-idealities of the embedded quantizer, associated either to its ADC section or to its DAC section – see for instance Fig. 2.19 and (2.35).
- Additional error power terms due to thermal noise, incomplete settling, time uncertainty (jitter), component non-linearities, charge feedthrough, . . .

Here we do not aim covering, even mentioning, all possible errors. However, readers must be aware that proper modeling of circuit errors is cornerstone to design robust, high-performance $\Sigma\Delta$ M ICs; specially whenever high-frequency, high-resolution and/or low-power are design targets.

In many practical cases, non-ideal error power contributions dominate. An example of such dominance is illustrated in Table 2.11. It summarizes contributions of a $\Sigma\Delta$ M designed for the analog front-end of an ADSL MoDem IC in a $0.25\mu\text{m}$ CMOS technology [Rio02b]. Note that the thermal noise terms are larger than the ideal quantization noise.

Table 2.11: Summary of error terms for an ADSL $\Sigma\Delta$ M in CMOS $0.25\mu\text{m}$ technology.

Quantization Noise	-88.1dB
Ideal quantization noise	-90.3dB
Amplifier DC-gain leak	-99.8dB
Capacitor mismatching leak ($\sigma = 0,1\%$ for 1pF)	-95.4dB
DAC non-linearity error	-96.4dB
Thermal Noise	-84.8dB
kT/C noise	-88.1dB
Opamp noise	-87.5dB
Clock Jitter	-90.1dB
In-Band Error Power	-82.3dB

In a situation like that in Table 2.11, the shaped quantization error becomes masked by other noise contributions. Such a type of masking is illustrated in Fig. 2.33 for a BP $\Sigma\Delta M$. The thick, dark line in this figure corresponds to the noise contributed by jitter (time uncertainty of the clock) ^{†24}. It is seen that this contribution is larger than that the ideal one obtained by just shaping the quantization.

Significant circuit non-idealities degrading the performance of DT $\Sigma\Delta M$ s include:

- Integrator leakage (finite DC gain).
- Analog coefficient tolerances due either to capacitor mismatch or to transistor mismatch; specially significant for cascade architectures since it precludes complete annulment of the quantization errors of stages different from the last one.
- ADC and DAC quantizer errors.
- Incomplete settling of the integrators; specially significant for high-frequency implementations.
- Jitter noise.
- Thermal noise and other sources of circuit noise.
- Nonlinear ON-resistance of the front-end switches.
- Charge injection and clock feedthrough.
- Voltage reference errors.

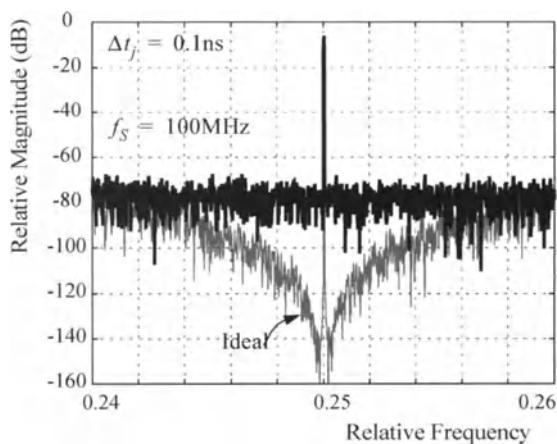


Figure 2.33: Illustrating noise-shaping masking caused by jitter.

24. In DT $\Sigma\Delta M$ jitter produces an in-band error power given by,

$$P_{\epsilon_{jitter}} \propto f^2 \cdot \frac{\sigma_{\Delta T}^2}{M}$$

whose inverse dependence with M is much slower than that of the ideal quantization noise. The effect of jitter is more pronounced for CT $\Sigma\Delta M$ s because it manifests as errors in the feedback signal $y(t)$ which are not shaped at all (see Fig. 2.31). It can be attenuated by properly choosing the DAC shape, which defines an important research area nowadays.

- Distortion produced by the components used for analog scaling (capacitors, transistors, resistors, . . .), by the opamps, . . .

Similar phenomena degrade the operation of CT $\Sigma\Delta$ Ms where errors associated to the DAC, and non-linearities associated to the front-end adder are of primary importance.

The first step towards coping with non-idealities is formulating their influence on the in-band error power. This results into a set of expressions,

$$P_{\varepsilon k} = P_{\varepsilon k}(\tilde{\Theta}, M, L) \quad \forall k \quad (2.42)$$

where k is an index spanning the set of relevant no-ideal phenomena, $\tilde{\Theta}$ is a vector of parameters describing the non-idealities, and functional dependences represented by $P_{\varepsilon k}(\bullet)$ change with the actual $\Sigma\Delta$ M architecture. Once these expressions are found, they must be solved to obtain the actual set of parameter values $\tilde{\Theta}_0$ meeting the design targets (resolution, bandwidth, power consumption, . . .). This is difficult process involving the solution of different trade-offs. Readers are referred to [Mede99a] for a comprehensive exposition of the problem supported by tailored methodologies and tools. Of course, after parameter values are found they have to be mapped onto adequate circuit topologies with properly chosen transistor and passive component sizes.

2.8 CONCLUSIONS

Coming back to Fig. 2.2, CMOS $\Sigma\Delta$ M ICs have been demonstrated for very large regions of the resolution versus conversion rate plane. It is demonstrated through the set of specifications listed in Table 2.1, Table 2.2, Table 2.3, Table 2.4, Table 2.6, Table 2.7, Table 2.9 and Table 2.10. Today $\Sigma\Delta$ M ICs still define a very active area of research with significant efforts being focused on devising architectures, error correction/calibration techniques, methodologies and circuits to make the oversampling ratio decrease.

2.9 REFERENCES

2.9.1 General References

- [Aziz93] P. Aziz, H.V. Sorensen, and J.V.D. Spiegel, “Multi-band Sigma-delta Modulation”. *IEE Electronics Letters*, pp. 760-762, April 1993.
- [Bair95] R.T. Baird and T. Fiez, “Linearity Enhancement of Multibit $\Delta\Sigma$ A/D and D/A Converters using Data Weighted Averaging”. *IEEE Trans. on Circuits*

- and Systems - Part II*, Vol. 42, pp. 753-762, December 1995.
- [Bree01] L. Breems and J.H. Huijsing, *Continuous-Time Sigma-Delta Modulation for A/D Conversion in Radio Receivers*. Kluwer Academic Publishers, 2001.
- [Broo02] T. Brooks, “Architecture Considerations for Multi-Bit $\Sigma\Delta$ ADCs”. in *Analog Circuit Design* (edited by M. Steyaert, A. Van Roermond and J.H. Huijsing). Kluwer Academic Publishers 2002.
- [Casi01] H.J. Casier, “Requirements for Embedded Converters in ADSL Communication Systems”. *Proc. of the 2001 IEEE ICECS*, Vol. 1, pp 489-492, September 2001.
- [Cher00] J.A. Cherry and W.M. Snelgrove, *Continuous-Time Delta-Sigma Modulators for High-Speed A/D Conversion*. Kluwer Academic Publishers, Boston 2000.
- [Corn94] R.F. Cornier Jr., T.L. Sculley, and R.H. Bamberger, “Combining Sub-band Decomposition and Sigma-delta Modulation for Wide-band A/D Conversion”. *Proc. of the 1994 IEEE Int. Symp. Circuits and Systems*, pp. 357-360, June 1994.
- [Enge99] J.V. Engelen and R. van de Plassche, *BandPass Sigma-Delta Modulators: Stability Analysis, Performance and Design Aspects*. Kluwer Academic Publishers, Boston 1999.
- [Eshr96] A. Eshragi and T. Fiez, “A Comparison of Three Parallel SD A/D Converters”. *Proc. of the 1996 IEEE Int. Symp. Circuits and Systems*, pp. 517-520, 1996.
- [Galt95] I. Galton and H.T. Jensen, “Delta-Sigma Based A/D Conversion Without Oversampling”. *IEEE Trans. Circuits and Systems - II*, Vol. 42, pp. 773-784, Dec. 1995.
- [Galt96] I. Galton and H.T. Jensen, “Oversampling Parallel Delta-Sigma Modulator A/D Conversion”. *IEEE Trans. Circuits and Systems - II*, Vol. 43, pp. 801-810, Dec. 1996.
- [Galt02] I. Galton, “Delta-Sigma Data Conversion in Wireless Transceivers”. *IEEE Trans. on Microwave Theory and Techniques*, Vol. 50, pp. 302-315, January 2002.
- [Khoi97] R. Khoini-Poorfard, L.B. Lim, and D.A. Johns, “Time-Interleaved Oversampling A/D Converters: Theory and Practice”. *IEEE Trans. Circuits and Systems - II*, Vol. 44, pp. 634-645, Aug. 1997.
- [King98] E.T. King, A. Eshragi, I. Galton, and T.S. Fiez, “A Nyquist-Rate Delta-Sigma A/D Converter”. *IEEE Journal of Solid-State Circuits*, Vol. 33, pp. 45-52, Jan. 1998.
- [Kiss00] P. Kiss et al., “Adaptive Digital Correction of Analog Errors in MASH ADC’s - Part II: Correction Using Test-Signal Injection”. *IEEE Trans. on Circuits and Systems - II*, Vol. 47, pp. 629-638, July 2000.
- [Koza00] M. Kozak and I. Kale, “Novel Topologies for Time-Interleaved Delta-Sigma Modulators”. *IEEE Trans. Circuits and Systems - II*, Vol. 47, pp. 639-654, July 2000.
- [Lee87] W.L. Lee and C.G. Sodini, “A Topology for Higher Order Interpolative Coders”. *Proc. of the 1987 IEEE Int. Symp. on Circuits and Systems*, pp. 459-462, 1987.
- [Lesl90] T.C. Leslie and B. Singh, “An Improved Sigma-Delta Modular Architec-

- ture”. *Proc. of the 1990 IEEE Int. Symp. on Circuits and Systems*, pp. 372-375, May 1990.
- [Leun92] B. Leung and S. Sutarja, “Multibit S-D A/D Converter Incorporating a Novel Class of Dynamic Element Matching Techniques”. *IEEE Trans. on Circuits and Systems - Part II*, Vol. 39, pp. 35-51, January 1992.
 - [Marq98b] A.M. Marques, V. Peluso, M.S.J. Steyaert, and W. Sansen, “Optimal Parameters for $\Delta\Sigma$ Modulator Topologies”. *IEEE Trans. on Circuits and Systems*, Vol. 45, pp. 1232-1241, September 1998.
 - [Mede99a] F. Medeiro, B. Pérez-Verdú, and A. Rodríguez-Vázquez, *Top-Down Design of High-Performance Sigma-Delta Modulators*. Kluwer Academic Publishers, Boston 1999.
 - [Nors97] S.R. Norsworthy, R. Schreier, and G.C. Temes (Editors), *Delta-Sigma Data Converters: Theory, Design and Simulation*. IEEE Press, New York 1997.
 - [OpT93] F. Op'T Eynde and W. Sansen, *Analog Interfaces for Digital Signal Processing Systems*. Kluwer Academic Publishers 1993.
 - [Paul99] A. Paul, H. Lee, J. Goodrich, T.F. Alailima and D. Santiago, “A Nyquist-Rate Pipelined Oversampling A/D Converters”. *IEEE Journal of Solid-State Circuits*, Vol. 34, pp. 1777-1787, December 1999.
 - [Ribn91] D.B. Ribner, “A Comparison of Modulator Networks for High-Order Oversampled $\Sigma\Delta$ Analog-to-Digital Converters”. *IEEE Transactions on Circuits and Systems*, Vol. 38, pp. 145-159, February 1991.
 - [Ri02a] R. del Río, F. Medeiro, J.M. de la Rosa, B. Pérez-Verdú, and A. Rodríguez-Vázquez, “Correction-Free Multi-Bit Sigma-Delta Modulators for ADSL”. in *Analog Circuit Design* (edited by M. Steyaert, A. Van Roermond and J.H. Huijsing). Kluwer Academic Publishers 2002.
 - [Rosa02] J.M. de la Rosa, B. Pérez-Verdú, and A. Rodríguez-Vázquez, *Systematic Design of CMOS Switched-Current Bandpass Sigma-Delta Modulators for Digital Communication Chips*. Kluwer Academic Publishers, 2002.
 - [Shoa95] O. Shoaei, *Continuous-Time Delta-Sigma A/D Converters for High Speed Applications*. Ph.D. Dissertation, Carleton University, 1995.
 - [Schr93] R. Schreier, “An Empirical Study of Higher Order Single Bit Sigma Delta Modulators”. *IEEE Trans. on Circuits and Systems - II*, Vol. 40, pp. 461-466, 1993.
 - [Wang00] X. Wang, W. Quin, and X. Ling, “Cascaded Parallel Oversampling Sigma-Delta Modulators”. *IEEE Trans. Circuits and Systems - II*, Vol. 47, pp. 156-161, Feb. 2000.

2.9.2 DT LP $\Sigma\Delta$ ICs

- [Adam86] R.W. Adams, “Design and Implementation of an Audio 18-bit Analog-to-Digital Converter Using Oversampling Techniques”. *Journal of Audio Engineering Society*, Vol. 34, pp. 153-166. March 1986.
- [Au97] S. Au and B.H. Leung, “A 1.95-V, 0.34-mW, 12-b Sigma-Delta Modulator Stabilized by Local Feedback Loops”. *IEEE Journal of Solid-State Circuits*, Vol. 32, March 1997.
- [Bajd02] O. Bajdechi and J.H. Huijsing, “A 1.8-V Delta-Sigma Modulator Interface

- for an Electret Microphone with On-Chip Reference". *IEEE Journal of Solid-State Circuits*, Vol. 37, pp. 279-285, March 2002.
- [Bair96] R.T. Baird and T.S. Fiez, "A Low Oversampling Ratio 14-b 500-kHz DS ADC with a Self-Calibrated Multibit DAC". *IEEE Journal of Solid-State Circuits*, Vol. 31, pp. 312-320, March 1996.
- [Bertl93] A. Bertl, H. Leopold, P. O'leary, and G. Winkler, "A Monolithic ADC for instrumentation Applications using Sigma-Delta Techniques". *Proc. of the 11th European Conf. on Circuit Theory and Design*, Vol. 2, pp. 1619-1624, 1993.
- [Bose88] B.E. Boser and B.A. Wooley, "The Design of Sigma-Delta Modulation Analog-to-Digital Converters". *IEEE Journal of Solid-State Circuits*, Vol. 23, pp. 1298-1308, December 1988.
- [Bran91a] B. Brandt and B.A. Wooley, "A 50-MHz Multibit $\Sigma\Delta$ Modulator for 12-b 2-MHz A/D Conversion". *IEEE Journal of Solid-State Circuits*, Vol. 26, pp. 1746-1756, December 1991.
- [Bran91b] B. Brandt, D.W. Wingard, and B.A. Wooley, "Second-Order Sigma-Delta Modulation for Digital-Audio Signal Acquisition". *IEEE Journal of Solid-State Circuits*, Vol. 23, pp. 618-627, April 1991.
- [Briga02] S. Brigati, F. Francesconi, P. Malcovati, and F. Maloberti, "A fourth-order single-bit switched-capacitor sigma-delta modulator for distributed sensor applications". *Proc. of the 2002 IEEE Instrumentation and Measurement Technology Conf.*, Vol.1, pp. 253-256, 2002.
- [Broo97] T.L. Brooks, D.H. Robertson, D.F. Kelly, A. Del Muro, and S. W. Harston, "A Cascaded Sigma-Delta Pipeline A/D Converter with 1.25 MHz Signal Bandwidth and 89 dB SNR". *IEEE Journal of Solid-State Circuits*, Vol. 32, pp. 1896-1906, December 1997.
- [Burm96] T.V. Burmas, K.C. Dyer, P.J. Hurst, and S.H. Lewis, "A second-order double-sampled delta-sigma modulator using additive-error switching". *IEEE Journal of Solid-State Circuits*, Vol. 31, pp. 284-293, March 1996.
- [Chen95] F. Chen and B.H. Leung, "A High Resolution Multibit Sigma-Delta Modulator with Individual Level Averaging". *IEEE Journal of Solid-State Circuits*, Vol. 30, pp. 453-460, April 1995.
- [Coba99] A.L. Coban and P.E. Allen, "A 1.5 V 1.0 mW Audio Delta Sigma Modulator with 98dB Dynamic Range". *Proc. of the 1999 IEEE Int. Solid-State Circuits Conf.*, pp. 50 -51, February 1999.
- [Dedi94] I. Dedic, "A Sixth-Order Triple-Loop $\Sigma\Delta$ CMOS ADC with 90dB SNR and 100kHz Bandwidth". *Proc. of the 1994 IEEE Int. Solid-State Circuits Conf.*, pp. 188-189, 1994.
- [Feld98] A.R. Feldman, B.E. Boser, and P.R. Gray, "A 13-Bit, 1.4-MS/s Sigma-Delta Modulator for RF Baseband Channel Applications". *IEEE Journal of Solid-State Circuits*, Vol. 33, pp. 1462-1469, October 1998.
- [Fogl00] E. Fogelman, I. Galton, W. Huff, and H. Jensen, "A 3.3-V single-poly CMOS audio ADC delta-sigma modulator with 538-dB peak SINAD and 105-dB peak SFDR". *IEEE Journal of Solid-State Circuits*, Vol. 35, pp. 297-307, March 2000.
- [Fogl01] E. Fogelman, J. Welz, and I. Galton, "An audio ADC Delta-Sigma modula-

- tor with 100-dB peak SINAD and 102-dB DR using a second-order mismatch-shaping DAC". *IEEE Journal of Solid-State Circuits*, Vol. 36, pp. 339-348, March 2001.
- [Fuji97] I. Fujimori et al., "A 5-V Single-Chip Delta-Sigma Audio A/D Converter with 111dB Dynamic Range". *IEEE Journal of Solid-State Circuits*, Vol. 32, pp. 329-336, March 1997.
- [Fuji00] I. Fujimori et al., "A 90dB SNR, 2.5MHz Output Rate ADC using Cascaded Multibit Delta-Sigma modulation at 8x Oversampling Ratio". *IEEE Journal of Solid-State Circuits*, Vol. 35, pp. 1820-1828, December 2000.
- [Geer99] Y. Geerts, A. Marques, M. Steyaert, and W. Sansen, "A 3.3-V, 15-bit, Delta-Sigma ADC with a Signal Bandwidth of 1.1 MHz for ADSL Applications". *IEEE Journal of Solid-State Circuits*, Vol. 34, pp. 927-936, July 1999.
- [Geer00] Y. Geerts, M. Steyaert, and W. Sansen, "A High Performance Multibit DS CMOS ADC". *IEEE Journal of Solid-State Circuits*, Vol. 35, pp. 1829-1840, December 2000.
- [Gome00] G.J. Gómez, "A 102-dB Spurious-Free DR Sigma-Delta ADC using a Dynamic Dither Scheme". *IEEE Transactions on Circuits and Systems-II*, pp. 531-535, June 2000.
- [Grilo96] J. Grilo, E. Mac Robbie, R. Halim, and G. Temes, "A 1.8V 94dB DR $\Sigma\Delta$ Modulator for Voice Applications". *Proc. of IEEE Int. Solid-State Circuits Conf.*, pp. 230-233, 1996.
- [Grilo02] J. Grilo, I. Galton, K. Wang, and R.G. Montemayor, "12-mW ADC Delta-Sigma Modulator with 80dB of Dynamic Range Integrated in a Single-Chip Bluetooth Transceiver". *IEEE Journal of Solid-State Circuits*, Vol. 37, pp. 271-278, March 2002.
- [Gupta02] S.K. Gupta and V. Fong, "A 64-MHz Clock-Rate SD ADC With 88-dB SNDR and -105-dB IM3 Distortion at a 1.5-MHz Signal Frequency". *IEEE Journal of Solid-State Circuits*, Vol. 37, pp. 1653-1661, December 2002.
- [Kash99] D.B. Kasha, W.L. Lee, and A. Thomsen, "A 16-mW, 120-dB linear switched-capacitor delta-sigma modulator with dynamic biasing". *IEEE Journal of Solid-State Circuits*, Vol. 34, pp. 921-926, July 1999.
- [Kert94] D.A. Kerth et al., "A 126 dB linear switched-capacitor delta-sigma modulator". *Proc. of the 1994 IEEE Int. Solid-State Circuits Conf.*, pp.196 -197, 1994.
- [King98] E.T. King, A. Eshragi, I. Galton, and T.S. Fiez, "A Nyquist-Rate Delta-Sigma A/D Converter". *IEEE Journal of Solid-State Circuits*, Vol. 33, pp. 45-52, Jan. 1998.
- [Kuo02] T.-H. Kuo, K.-D. Chen, and H.-R. Yeng, "A Wideband CMOS Sigma-Delta Modulator With Incremental Data Weighted Averaging". *IEEE Journal of Solid-State Circuits*, Vol. 37, pp. 11-17, January 2002.
- [Lamp01] H. Lampinen and O. Vainio, "Low-voltage fourth-order CMOS sigma-delta modulator implementation". *Electronics Letters*, Vol. 37, pp. 734-735, June 2001.
- [Leun97] K.Y. Leung et al., "A 5V, 118dB $\Delta\Sigma$ Analog-to-Digital Converter for Wideband Digital Audio". *Proc. of IEEE Int. Solid-State Circuit Conf.*, pp. 218-219, 1997.

- [Marq98a] A.M. Marques, V. Peluso, M.S.J. Steyaert, and W. Sansen, "A 15-b Resolution 2-MHz Nyquist Rate $\Delta\Sigma$ ADC in a 1- μm CMOS Technology". *IEEE Journal of Solid-State Circuits*, Vol. 33, pp. 1065-1075, July 1998.
- [Maul00] P.C. Maulik, M.S. Chadha, W.L. Lee, and P.J. Crawley, "A 16-bit 250-kHz delta-sigma modulator and decimation filter". *IEEE Journal of Solid-State Circuits*, Vol. 35, pp. 458-467, April 2000.
- [Mede97] F. Medeiro, B. Pérez-Verdú, José M. de la Rosa, and A. Rodríguez-Vázquez, "Using CAD Tools for Shortening the Design Cycle of High-Performance $\Sigma\Delta$ M: A 16.4bit 1.71mW $\Sigma\Delta$ M in CMOS 0.7 μm Technology". *International Journal of Circuit Theory and Applications*, Vol. 25, pp. 319-334, 1997.
- [Mede99b] F. Medeiro, B. Perez-Verdú, and A. Rodríguez-Vázquez, "A 13-bit, 2.2-MS/s, 55-mW Multibit Cascade $\Sigma\Delta$ Modulator in CMOS 0.7- μm Single-Poly Technology". *IEEE Journal of Solid-State Circuits*, Vol. 34, n. 6, pp. 748-760, June 1999.
- [Miao98] G. Miao, H.C. Yang, and T. Pushan, "An oversampled A/D converter using cascaded fourth order sigma-delta modulation and current steering logic". *Proc. of the 1998 IEEE Int. Symp. on Circuits and Systems*, Vol. 1, pp. 412-415, 1998.
- [Mille02] M.R. Miller and C.S.A. Petrie, "A multi-bit sigma-delta ADC for multi-mode receivers". *Proc. of the 2002 IEEE Custom Integrated Circuits Conf.*, pp. 191-194, 2002.
- [Mori00] J.C. Morizio et al., "14-bit 2.2-MS/s Sigma-Delta ADC's". *IEEE Journal of Solid-State Circuits*, Vol. 35, pp. 968-976, July 2000.
- [Nade94] S. Nadeem, C.G. Sodini, and L. Hae-Seung, "16-channel oversampled analog-to-digital converter". *IEEE Journal of Solid-State Circuits*, Vol. 29, pp. 1077-1085, Sept. 1994.
- [Nys97] O. Nys and R. Henderson, "A 19-bit Low-Power Multi-Bit Sigma-Delta ADC Based on Data Weighted Averaging". *IEEE Journal of Solid-State Circuits*, Vol. 32, pp. 933-942, July 1997.
- [Olia02] O. Oliae, P. Clement, and P. Gorisse, "A 5-mW sigma-delta modulator with 84-dB dynamic range for GSM/EDGE". *IEEE Journal of Solid-State Circuits*, Vol. 37, pp. 2-10, Jan. 2002.
- [Pelu96] V. Peluso, M. Steyaert, and W. Sansen, "A Switched Opamp 1.5V - 100 μW $\Sigma\Delta$ Modulator with 12 bits Dynamic Range". *Proc. of the 1996 European Solid-State Circuit Conf.*, pp. 256-259, September 1996.
- [Pelu98] V. Peluso, P. Vancorenland, A.M. Marques, M.S.J. Steyaert, and W. Sansen, "A 900mW Low-Power Delta-Sigma A/D Converter with 77dB". *IEEE Journal of Solid-State Circuits*, Vol. 33, pp. 1887-1897, Dec. 1998.
- [Rabi97] S. Rabii and B.A. Wooley, "A 1.8V Digital-Audio Sigma-Delta Modulator in 0.8 μm CMOS". *IEEE Journal of Solid-State Circuits*, Vol. 32, pp. 783-796, June 1997.
- [Rebe90] M. Rebeschini et al., "A 16-b 160 kHz CMOS A/D Converter using Sigma-Delta Modulation". *IEEE Journal of Solid-State Circuits*, Vol. 25, pp. 431-440, April 1990.
- [Rio01] R. del Río, J.M. de la Rosa, F. Medeiro, B. Pérez-Verdú, and A. Rodríguez-Vázquez, "High-performance sigma-delta ADC for ADSL applications in

- 0.35 μ m CMOS digital technology”. *Proc. of the 2001 IEEE Int. Conf. on Electronics, Circuits and Systems*, Vol. 1, pp. 501-504, August 2001.
- [Rio02b] R. del Río, F. Medeiro, J.M. de la Rosa, B. Pérez-Verdú, and A. Rodríguez-Vázquez, “A 2.5-V Sigma-Delta modulator in 0.25 μ m CMOS for ADSL”. *Proc. of the 2002 IEEE Int. Symp. on Circuits and Systems*, Vol. 3, pp. 301-304, 2002.
- [Rito94] T. Ritoniemi, E. Pajarre, S. Ingalsuo, T. Husu, V. Eerola, and T. Saramiki, “A stereo audio sigma-delta A/D-converter”. *IEEE Journal of Solid-State Circuits*, Vol. 29, pp. 1514-1523, Dec. 1994.
- [Romb01] P. Rombouts, W. de Wilde, and L. Weyten, “A 13.5-b 1.2-V Micropower Extended Counting A/D Converter”. *IEEE Journal of Solid-State Circuits*, Vol. 36, pp. 176-183, February 2001.
- [Sarh93] M. Sarhang-Nejad and G.C. Temes, “A High-Resolution $\Sigma\Delta$ ADC with Digital Correction and Relaxed Amplifiers Requirements”. *IEEE Journal of Solid-State Circuits*, Vol. 28, pp. 648-660, June 1993.
- [Send97] D. Senderowicz et al., “Low-Voltage Double-Sampled $\Sigma\Delta$ Converters”. *Proc. of 1997 IEEE Int. Solid-State Circuits Conf.*, pp. 210-211, 1997.
- [Snoe01] M.F. Snoeij, O. Bajdechi, and J.H. Huijsing, “A 4th-order switched-capacitor sigma-delta A/D converter using a high-ripple Chebyshev loop filter”. *Proc. of the 2001 IEEE Int. Symp. on Circuits and Systems*, Vol. 1, pp. 615-618, 2001.
- [Than97] C.K. Thanh, S.H. Lewis, and P.J. Hurst, “A second-order double-sampled delta-sigma modulator using individual-level averaging”. *IEEE Journal of Solid-State Circuits*, Vol. 32, pp. 1269-1273, August 1997.
- [Tili01] T. Tilie, J. Sauerbrey, and D. Schmitt-Landsiedel, “A 1.8-V MOSFET-only Sigma-Delta modulator using substrate biased depletion-mode MOS capacitors in series compensation”. *IEEE Journal of Solid-State Circuits*, Vol. 36, pp. 1041-1047, July 2001.
- [Vleu01] K. Vleugels, S. Rabii, and B. Wooley, “A 2.5V Broadband Multi-Bit $\Sigma\Delta$ Modulator with 95dB Dynamic Range”. *Proc. of the 2001 IEEE Int. Solid-State Circuit Conf.*, pp. 50-51, Feb. 2001.
- [Wang01] C.B. Wang, “A 20-bit 25-kHz delta-sigma A/D converter utilizing a frequency-shaped chopper stabilization scheme”. *IEEE Journal of Solid-State Circuits*, Vol. 36, pp. 566-569, March 2001.
- [Will94] L.A. Williams III and B.A. Wooley, “A Third-Order Sigma-Delta Modulator with Extended Dynamic Range”. *IEEE Journal of Solid-State Circuits*, Vol. 29, pp. 193-202, March 1994.
- [Yama94] K. Yamamura, A. Nogi, and A. Barlow, “A low power 20 bit instrumentation delta-sigma ADC”. *Proc. of the 1994 IEEE Custom Integrated Circuits Conf.*, pp. 519 -522, 1994.
- [Yin93] G.M. Yin, F. Stubbe, and W. Sansen, “A 16-bit 320kHz CMOS A/D Converter using 2-Stage 3rd-Order $\Sigma\Delta$ Noise-Shaping”. *IEEE Journal of Solid-State Circuits*, Vol. 28, pp. 640-647, June 1993.
- [Yin94] G. Yin and W. Sansen, “A High-Frequency and High-Resolution Fourth-Order $\Sigma\Delta$ A/D Converter in BiCMOS Technology”. *IEEE Journal of Solid-State Circuits*, Vol. 29, pp. 857-865, August 1994.

- [Yoon98] S.H. Yoon and H.S.M. Rezaul “A sixth-order CMOS sigma-delta modulator”. *Proc. of the 1998 IEEE Int. ASIC Conference*, pp. 63 -66, 1998.

2.9.3 BP Sigma-Delta Modulator ICs

- [Andr96] E. André, D. Morche, F. Balestro, and P. Senn, “A 2-path $\Sigma\Delta$ Modulator for Bandpass Application”. *Proc. of the 22nd European Solid-State Circuits Conf.*, pp. 216-219, September 1996.
- [Andr98] E. André, G. Martel, D. Morche, and P. Senn, “A Bandpass A/D Converter and its I/Q Demodulator for Integrated Receivers”. *Proc. of the 24th European Solid-State Circuits Conf.*, pp. 416-419, September 1998.
- [Baza98] S. Bazarjani and M. Snelgrove, “A 160MHz Fourth-Order Double-Sampled SC Bandpass Sigma-Delta Modulator”. *IEEE Trans. Circuits and Systems-II*, Vol. 45, pp. 547-555, May 1998.
- [Baza99] S. Bazarjani, S. Younis, J. Goldblatt, D. Butterfield, G. McAllister, and S. Ciccarelli, “An 85MHz IF Bandpass Sigma-Delta Modulator for CDMA Receivers”. *Proc. of the 25th European Solid-State Circuits Conf.*, pp. 266-269, September 1999.
- [Corm97] R.F. Cormier, T.L. Sculley, and R.H. Bamberger, “A Fourth Order Bandpass Delta-Sigma Modulator with Digitally Programmable Passband Frequency”. *Int. Journal of Analog Integrated Circuits and Signal Processing*, pp. 217-229, 1997.
- [Cheu01] V.S.L. Cheung, H.C. Luong, and W.H. Ki, “A 1V 10.7MHz Switched-Opamp Bandpass $\Sigma\Delta$ Modulator using Double-Sampling Finite-Gain Compensation Technique”. *Proc. of the 2001 IEEE Int. Solid-State Circuit Conf.*, pp. 52-53, February 2001.
- [Chua98] S. Chuang, H. Liu, X. Yu, T.L. Sculley, and R.H. Bamberger, “Design and Implementation of Bandpass Delta-Sigma Modulators Using Half-Delay Integrators”. *IEEE Trans. Circuits and Systems-II*, Vol. 45, pp. 535-546, May 1998.
- [Cusi00] P. Cusitano, F. Stefani, and A. Baschirotto, “A 73dB SFDR 10.7MHz 3.3V CMOS Bandpass $\Sigma\Delta$ Modulator sampled at 37.05MHz”. *Proc. of the 26th European Solid-State Circuits Conf.*, pp. 80-83, September 2000.
- [Haira96] A. Hairapetian, “An 81-MHz IF Receiver in CMOS”. *IEEE Journal of Solid-State Circuits*, pp. 1981-1986, December 1996.
- [Jant92] S.A. Jantzi, W.M. Snelgrove, and P.F. Ferguson, “A 4th-Order BandPass Sigma-Delta Modulator”. *Proc. of the 1992 Custom Integrated Circuit Conf.*, pp. 16.5.1-4, May 1992.
- [Jant93] S.A. Jantzi, M. Snelgrove, and P.F. Ferguson, “A Fourth-Order Bandpass Sigma-Delta Modulator”. *IEEE Journal of Solid-State Circuits*, Vol. 28, pp. 282-291, March 1993.
- [Jant97] S.A. Jantzi, K.W. Martin, and A.S. Sedra, “Quadrature Bandpass $\Delta\Sigma$ Modulation for Digital Radio”. *IEEE Journal of Solid-State Circuits*, Vol. 32, pp. 1935-1949, December 1997.
- [Liu97] H. Liu, X. Yu, T. Sculley, and R.H. Bamberger, “A Fourth Order Bandpass Delta-Sigma A/D Converter with Input Modulation Network and Digitally

- Programmable Passband”. *Proc. of the 1997 IEEE Int. Symp. Circuits and System*, pp. 385-388, May 1997.
- [Long93] L. Longo and H. Bor-Rong, “A 15b 30KHz Bandpass Sigma-Delta Modulator”. *Proc. of the 1993 IEEE Int. Solid-State Circuit Conf.*, pp. 226-227, March 1993.
- [Ong97] A.K. Ong and B.A. Wooley, “A Two-Path Bandpass $\Sigma\Delta$ Modulator for Digital IF Extraction at 20MHz”. *IEEE Journal of Solid-State Circuit*, Vol. 32, pp. 1920-1933, December 1997.
- [Park99] J. Park, E. Joe, M. Choe, and B.P. Song, “A 5-MHz IF Digital FM Demodulator”. *IEEE Journal of Solid-State Circuit*, Vol. 34, pp. 3-11, January 1999.
- [Rosa00] J.M. de la Rosa, B. Pérez-Verdú, R. del Río, and A. Rodríguez-Vázquez, “A CMOS $0.8\mu\text{m}$ Transistor-Only 1.63MHz Switched-Current Bandpass $\Sigma\Delta$ Modulator for AM Signal A/D Conversion”. *IEEE Journal of Solid-State Circuits*, Vol.35, pp. 1220-1226, August 2000.
- [Salo02] T. Salo, S. Lindfors, and K. Halonen, “A 80-MHz Bandpass $\Sigma\Delta$ Modulator for a 100-MHz IF Receiver”. *IEEE Journal of Solid-State Circuits*, Vol. 37, pp. 798-808, July 2002.
- [Song95] B.S. Song, “A Fourth-Order Bandpass Delta-Sigma Modulator with Reduced Number of Op Amps”. *IEEE Journal of Solid-State Circuits*, Vol. 25, pp. 1309-1315, December 1995.
- [Taba99] A. Tabatabaei and B. Wooley, “A Wideband Bandpass Sigma-Delta Modulator for Wireless Applications”. *Proc. of the 1999 IEEE Int. Symp. on VLSI Circuits*, pp. 91-92, 1999.
- [Taba00] A. Tabatabaei and B. Wooley, “A Two-path Bandpass Sigma-Delta Modulator with Extended Noise Shaping”. *IEEE Journal Solid-State Circuits*, Vol. 35, pp. 1799-1809, December 2000.
- [Tao99] H. Tao and J.M. Khouri, “A 400-MS/s Frequency Translating Bandpass Sigma-Delta Modulator”. *IEEE Journal of Solid-State Circuits*, Vol. 34, pp. 1741-1752, December 1999.
- [Toni99] D. Tonietto, P. Cusinato, F. Stefani, and A. Baschirrotto, “A 3.3V CMOS 10.7MHz 6th-order bandpass $\Sigma\Delta$ Modulator with 78dB Dynamic Range”. *Proc. of the 1999 European Solid-State Circuits Conf.*, pp. 78-81, September 1999.
- [Ueno02] T. Ueno, A. Yasuda, T. Yamaji, and T. Itakura, “A Fourth-Order Bandpass $\Delta-\Sigma$ Modulator Using Second-Order Bandpass Noise-Shaping Dynamic Element Matching”. *IEEE Journal of Solid-State Circuits*, Vol. 37, pp. 809-816, July 2002.

2.9.4 CT $\Sigma\Delta$ ICs

- [Abou01] H. Aboushady and M.-M. Louerat, “Systematic Design of High-Linearity Current-Mode Integrators for Low-Power Continuous-Time Sigma-Delta Modulators”. *Proc. of the 8th IEEE Int. Conf. on Electronics, Circuits and Systems*, Vol. 2, pp. 963-966, September 2001.
- [Bree00] L.J. Breems, E.J. Van der Zwan and J. Huijsing, “A 1.8-mW CMOS $\Sigma\Delta$ Modulator with Integrated Mixer for A/D Conversion of IF Signals”. *IEEE*

- Journal of Solid-State Circuits*, Vol. 35, pp. 468-475, April 2000.
- [Enge99] J.A.E.P. van Engelen, R.J. van de Plassche, E. Stikvoort, and A.G. Venes, "A Sixth-Order Continuous-Time Bandpass Sigma-Delta Modulator for Digital Radio IF". *IEEE Journal of Solid-State Circuits*, Vol. 34, pp. 1753-1764, Dec. 1999.
- [Lin99] C.H. Lin, C. Shi, M. Ismail, and Moon Gyu. "A 5 MHz Nyquist Rate Continuous-Time Sigma-Delta Modulator for Wideband Wireless Communication". *Proc. of the 1999 IEEE Int. Symp. on Circuits and Systems*, Vol. 2, pp. 368-371, May 1999.
- [Gerf01] F. Gerfers, M. Ortmanns and Y. Manoli, "A 12-bit Power Efficient Continuous-Time $\Sigma\Delta$ Modulator with 220 μ W Power Consumption". *Proc. of the 2001 European Solid-State Circuits Conference*, pp. 536-539, September 2001.
- [Gerf02] F. Gerfers, M. Ortmanns, L. Samid and Y. Manoli, "Implementation of a 1.5V Low-Power Clock-Jitter Insensitive Continuous-Time Sigma-Delta Modulator". *Proc. of the 2002 IEEE Int. Symp. on Circuits and Systems*, Vol. 2, pp. 652-655, May 2002.
- [Hall92] B. Hallgren, "Design of a Second-Order CMOS Sigma-Delta A/D Converter with 150MHz Clock Rate". *Proc. of the 1992 European Solid-State Circuits Conf.*, pp. 103-106, September 1992.
- [Hall93] B. Hallgren, "Continuous-Time Implementation of Sigma-Delta Converters for High Frequency Operation". *Proc. of the 1993 European Conference on Circuit Theory and Design*, pp. 1601-1606, September 1993.
- [Hsu00] I. Hsu and H.C. Luong, "A 70-MHz Continuous-Time CMOS bandpass Sigma-Delta Modulator for GSM Receivers". *Proc. of the 2000 IEEE Int. Symp. on Circuits and Systems*, Vol. 3, pp. 750-753, May 2000.
- [Luh98a] L. Luh, J. Choma Jr., and J. Draper, "A 50-MHz Continuous-Time Switched-Current Sigma-Delta Modulator". *Proc. of the 1998 IEEE Int. Symp. on Circuits and Systems*, Vol. 1, pp. 579-582, 1998.
- [Luh98b] L. Luh and J. Choma, "A Continuous-Time Switched-Current Sigma-Delta Modulator with Reduced Loop Delay". *Proc. of the 8th Great Lakes Symp. on VLSI*, pp. 286-291, 1998.
- [Luh00] L. Luh, J. Choma and J. Draper, "A 400-MHz 5th-Order Continuous-Time Switched-Current $\Sigma\Delta$ Modulator". *Proc. of the 2000 European Conf. on Solid-State Circuits*, pp. 72-75, September 2000.
- [Redm94] W. Redman-White, A.M. Durham, "Integrated fourth-order SD convertor with stable self-tuning continuous-time noise shaper". *IEE Proc. Circuits Devices Syst.*, Vol. 141, pp. 145-150, June 1994.
- [Tao99] H. Tao and J.M. Khouri, "A 400MS/s Frequency Translating BandPass Delta-Sigma Modulator". *IEEE Journal of Solid-State Circuits*, Vol. 34, pp. 1741-1752, December 1999.
- [Veld02] R.H.M. van Veldhoven, B.J. Minnis, H.A. Hegd and A.H.M. Roermund, "A 3.3-mW $\Sigma\Delta$ Modulator for UMTS in 0.18- μ m CMOS With 70-dB Dynamic Range in 2-MHz Bandwidth". *IEEE Journal of Solid-State Circuits*, Vol. 37, pp. 1645-1652, December 2002.
- [Zwan96] E.J. van der Zwan and E.C. Dijkmans, "A 0.2mW CMOS $\Sigma\Delta$ Modulator for

- Speech Coding with 80dB Dynamic Range”. *IEEE Journal of Solid-State Circuits*, pp. 1873-1880, December 1996.
- [Zwan97] E.J. van der Zwan, “A 2.3mW CMOS $\Sigma\Delta$ Modulator for Audio Applications”. *Proc. of the 1997 IEEE Int. Solid-State Circuits Conf.*, pp. 220-221, 1997.
- [Zwan99] E.J. van der Zwan et al., “A 13mW 500kHz Data Acquisition IC with 4.5 Digit DC and 0.02% Accurate True-RMS Extraction”. *Proc. of the 1999 IEEE Int. Solid-State Circuits Conf.*, pp. 398-399, 1999.
- [Zwan00] E.J. van der Zwan and C.A.A. Bastiaansen, “A 10.7-MHz IF-to-Baseband SD A/D Conversion System for AM/FM Radio Receivers”. *IEEE Journal of Solid-State Circuits*, Vol. 35, pp. 1810-1819, December 2000.

Chapter 3

Current-Steering High-Speed D/A Converters for Communications

José Luis González, and Eduard Alarcón

3.1 INTRODUCTION AND STATE OF THE ART

Modern broadband communication integrated circuits require as fundamental subcircuits digital to analog converters (DAC) exhibiting both high-speed and high-resolution [1]. Wide bit-count DACs working at sampling clock frequencies in the range of the hundred of MHz will continue to be required, hence dictating Nyquist-rate data conversion, as for instance to convert digital bitstreams into continuous-time signals prior to up-conversion mixers preceding RF transmitters in wireless systems or to drive digital cable communications modems. This high rates have also been hitherto required by high-resolution displays for computer graphics and modern HDTV systems, for which time-domain performance is of utmost relevance, although spectral performance demands are much more stringent for communication ICs in which DACs are used to synthesize complex waveforms for which frequency-domain high-performance has to be attained [2]. The development of future mobile communication systems (including both 3G terminals and basestations) as well as the prospective use of ubiquitous communication systems will continue the trust towards high-performance DAC conversion stages.

This chapter provides a general vision on Nyquist-rate current steering DACs, since these architectures are intrinsically suited, as compared to their counterparts based on resistor or capacitor ladders, to both high speed, low power consumption and integrability in standard digital CMOS technologies (hence

being suited to be embedded in a system-on-a-chip approach yielding low cost and high reliability). In particular, the spectral performance of the DAC circuit is emphasized throughout the chapter.

The chapter is structured as follows. Section 4.1 continues after the introduction with a brief review of commonly used architectural options for DAC circuits, discussing the advantages and disadvantages of each approach. The current steering topology, in particular that of segmented architecture, is highlighted as a proper candidate for the performance level required in high-speed low-power communication systems. Additionally, a performance comparison for several actual current-steering DAC implementations is presented so as to provide a review of the state of the art in this field. Section 4.2 is devoted to the theoretical study of the various error sources specific to current-steering DACs, both mismatch and non-ideal output impedance of the basic current cell, as well as their impact on both static and spectral performance metrics, such as INL, DNL and SFDR. The purpose of section 4.3 is to present the basic circuit-level implementation issues of a current-steering DAC, which allows to discuss effects such as technology-related mismatch and finite output impedance as well as the nature of dynamic errors such as the behavior of switch transistors, glitches, settling time and clock jitter, while studying their effects on the spectral performance of the DAC. Section 4.4 covers several architecture-level methods that intend to diminish systematic errors in the current cell matrix implementation. Both one-dimensional and two-dimensional switching sequence design techniques are discussed and complemented with other techniques. The target of section 4.5 is to provide criteria for obtaining the optimum segmentation degree in segmented architectures, either based on an area *vs* static performance trade-off or on an spectral performance based optimum segmentation. The chapter concludes in section 4.6 with a compendium of practical design aspects, both a circuit sizing procedure which takes into account the complete design space, and layout issues, which are finally illustrated in a design example.

3.1.1 Architectural alternatives for Nyquist rate D/A converters

On an historical timescale, high-speed converters were originally implemented as standalone integrated circuits using fast bipolar technologies. Driven afterwards by system-level specifications, low-performance CMOS DACs were followed by high-performance CMOS DACs for time-domain applications (mostly in the video systems area) up to the CMOS DACs for frequency-domain applications currently used in communication systems and capable of being embedded in complex systems-on-a-chip (SoC).

Conceptually, DAC converters can be initially classified into high-speed and low-speed architectures. The latter (including serial DACs and algorithmic architectures with the potential use of pipelining) are not discussed in this chapter, since they are seldom applied in communications. On the other hand,

high-speed DACs may operate either with Nyquist rate (which is strictly required for high-bandwidth applications) as opposite to DACs with oversampling (*e.g.* for audio or video applications). The study of the former, which are theoretically able to operate within the whole frequency range according to the sampling theorem, is the purpose of this chapter.

More specifically, Nyquist rate high-speed DAC converters can be classified following different criteria. Depending on the nature of the scaling elements (and thus the magnitudes to be scaled) they can be divided into the so-called voltage-scaling (or voltage-mode), current-scaling (or current-mode) and charge-scaling or charge distribution DAC converters.

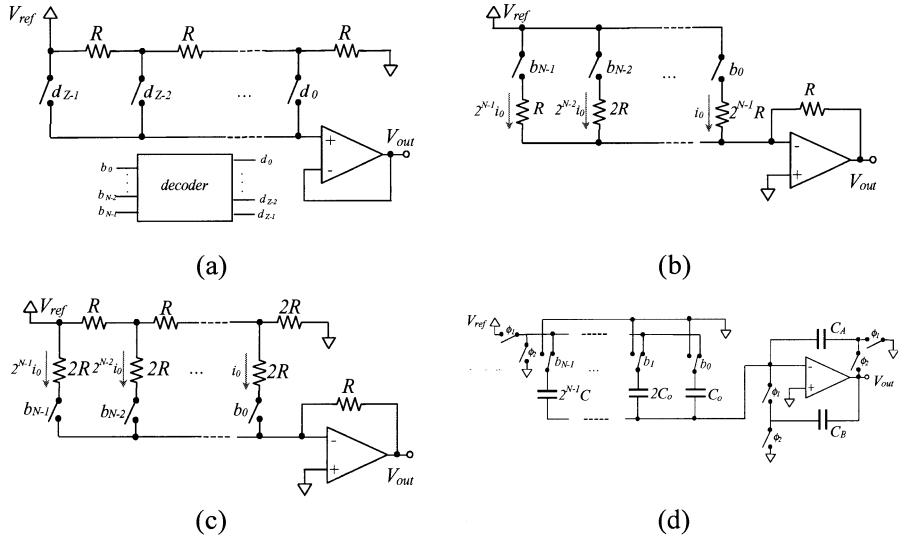


Figure 3.1. Architectural alternatives for Nyquist rate D/A converters (a) Voltage scaling (b) Current-scaling based on passive binary-weighted resistor ladder (c) Current-scaling based on passive R-2R ladder (d) Charge-scaling based on switched-capacitor circuit

As far as voltage-scaling DAC converters are concerned [3] –Figure 3.1 (a)–, they are based on the fact that the attenuation of a reference voltage can be obtained in a straightforward manner by a resistive divider composed of a string of equally-valued resistors, while the selection of the voltage corresponding to a particular digital code is carried out by means of a proper decoder (usually implemented with a tree of switches [4], which justifies their alternative denomination of encoded architectures; note that they can be interpreted as a digital potentiometer as well). It is immediate to demonstrate that the voltage across k resistors is given by:

$$V_k = \frac{kR}{2^N R} V_{ref}, \quad (3.1)$$

thus providing the functionality of the DAC conversion. Although one of the interesting advantages of this architecture is its inherent monotonicity, the structure lacks high-speed capabilities due to its sensitivity to parasitic capacitances related to the existence of high impedance nodes, requiring in addition a set of 2^N equally sized resistors.

Being the aim of a DAC converter to reproduce an analog output signal (either voltage or current) that is a representation of the input digital word, the rest of architectures pursue to implement the required functionality for this purpose, this is, to provide an output signal s_{out}

$$s_{out} = s_{ref} \left(b_0 2^{-1} + b_1 2^{-2} + \dots + b_{N-1} 2^{-N} \right) = s_{ref} \sum_{j=0}^{N-1} b_j 2^{-(j+1)} = s_{ref} k, \quad (3.2)$$

where s represents either a voltage or current signal, N the converter resolution in number of bits, and $0 < k < 2^N - 1$ stands for the normalized analog equivalent of the input digital word $D = [b_0, b_1, \dots, b_{N-1}]$. The step size Δ is defined as the analog value corresponding to one Least Significant Bit (LSB) of the digital code, and is related to the Full Scale range (FS) as $\Delta = FS/2^N$. Since the aggregation of signals is connatural to current representation and, in addition, current scaling is straightforward in that domain, the architectures based on current-scaling are a natural option. Figure 3.1 (b) corresponds to a resistor ladder responsible of providing a binary-weighted v/i conversion as to sum scaled currents. The architecture is scarcely used in practice because the high spread of resistance values is unpractical to accomplish, being actual implementations based on R-2R structures, as shown in Figure 3.2 (c). Both architectures have an historical origin and correspond mostly to bipolar processes, even though they have recently been reconsidered in CMOS environments by taking advantage of MOSFET-only ladders [5].

Voltage to charge conversion provided by scaled capacitors can be taken advantage of to derive a switched capacitor DAC structure, an illustration of which is depicted in Figure 3.1(d). Note that this architecture can use binary-weighted elements or a C/2C approach as well [3].

The previous current and charge scaling architectures can even be combined into an hybrid topology [3]. Additionally, cascading or semialgorithmic approaches can be considered [6].

Previous architectures share the disadvantage of requiring an output (transimpedance) amplifier, which apart from occupying a non-negligible extra area and represent extra power consumption (which becomes particularly important for high-rate DACs), notably reduces the linearity of the DAC. Therefore, for high-performance applications, the use of this buffer is

disregarded and a current-mode output with the capability of directly driving off-chip low-ohmic resistors is a must. This imposes restrictive conditions to the scaling elements as regards their impedance in order to avoid severe loading effects (even when driving low-impedance loads, as is the case of $50\ \Omega$ or $75\ \Omega$ of coaxial cables). The solution to this design problem considers active current cells as high-impedance current-scaling elements as opposite to passive elements. Figure 3.2 (a) shows a conceptual scheme of this architecture, in which it can be appreciated that the switching of the N binary-weighted scaled current cells naturally reproduces the ideal functionality given by eq. (3.2). Note that this topology exhibits excellent power efficiency since all the power of the circuit is directed to the output.

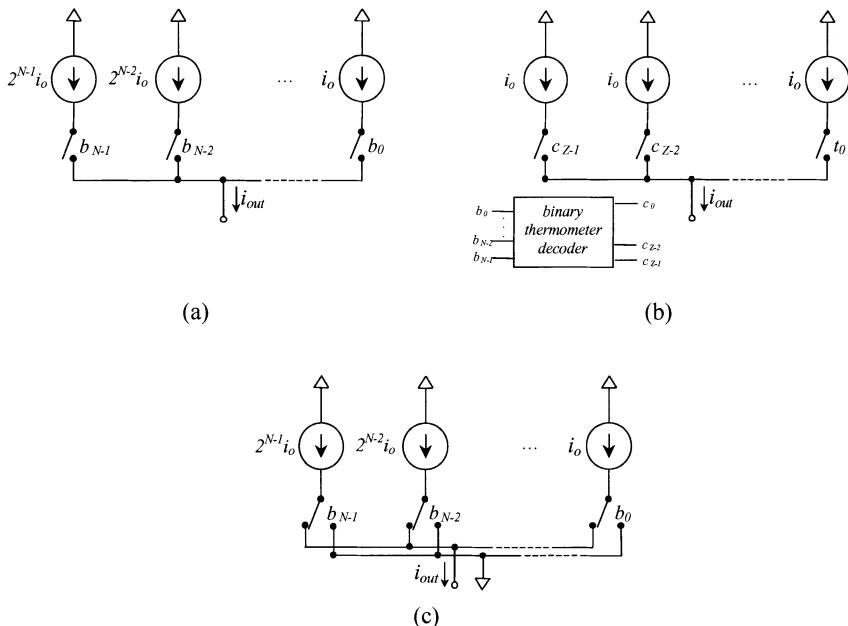


Figure 3.2. Architectural alternatives for Nyquist rate D/A converters. Current-scaling based on active current sources (a) Binary-weighted (b) Thermometer-coded (c) Current-steering approach

Although binary-weighted architectures require minimum number of elements (both passive or active scaling elements and switches), they suffer from high sensitivity to element mismatch. Hence, additional possible criteria of classification takes into account the nature of the digital code that switches the elements. DAC converters can be further divided into binary-weighted or thermometer-code converters. The latter approach takes advantage of a set of

equally sized elements and a thermometer code to derive the digital-to-analog characteristic:

$$s_{out} = s_{ref} (c_0 + c_1 + \dots + c_{2^N - 1}) = s_{ref} \sum_{j=0}^{2^N - 1} c_j. \quad (3.3)$$

The advantage of the thermometer code structure is its reduced sensitivity to element matching, being in particular monotonic in its input-output characteristic. On the other side, the number of scaling elements and switches grows exponentially with the number of bits, thus precluding its use for high resolutions. Figure 3.2 (b) depicts a thermometer code architecture for the current-scaling circuit. Thermometer-code operated versions of passive element architectures (both resistive and capacitive based) can also be envisaged.

Resorting to the active current scaling architectures in Figure 3.2 (a),(b), although the simplest means of activating a current source is to switch its output on/off, this solution is impractical because of current glitches. The problem is circumvented by using current steering, this is, to steer or redirect the current either to the output summing node or to a dummy low-impedance node by means of a complementary switch. By following this approach, which corresponds to the architecture of Figure 3.2 (c), the basic current cell is never left in open circuit.

Given the fact that the different architectures present both advantages and disadvantages, it is plausible to conclude that a combination of different approaches within a single hybrid architecture can result in a good trade-off candidate. This popular approach has the most notable representative for high-speed high-resolution applications in the current-steering segmented architecture, which is discussed in the subsequent subsection and constitutes the core of this chapter.

3.1.2 Current-steering D/A converter actual implementations

As shown in the segmented current-steering segmented DAC of Figure 3.3(a), the Least Significant Bits (LSB's) steer a binary weighted array of current sources –binary segment- whilst the Most Significant Bits (MSB's) are decoded thermometer-wise to steer a unary array of current sources –thermometer segment-.

Figure 3.3 (b) shows a practical block diagram of a N -bit current-steering segmented DAC which represents the common structure of current CMOS implementations (and even their floorplan). The B lowest input bits are the binary code segment of the converter, which after a latency equalizer directly control the switches of the B LSB binary weighted output current sources. The M highest bits are the thermometer code segment. They are decoded to drive the switching of 2^M output current sources, all of them of the same unary weight. The size of

the decoder increases exponentially with M . The update rate of the circuit is synchronously provided by a clock signal, which is distributed to the whole array via a clock driver. The output differential current is directly capable of driving low impedance loads.

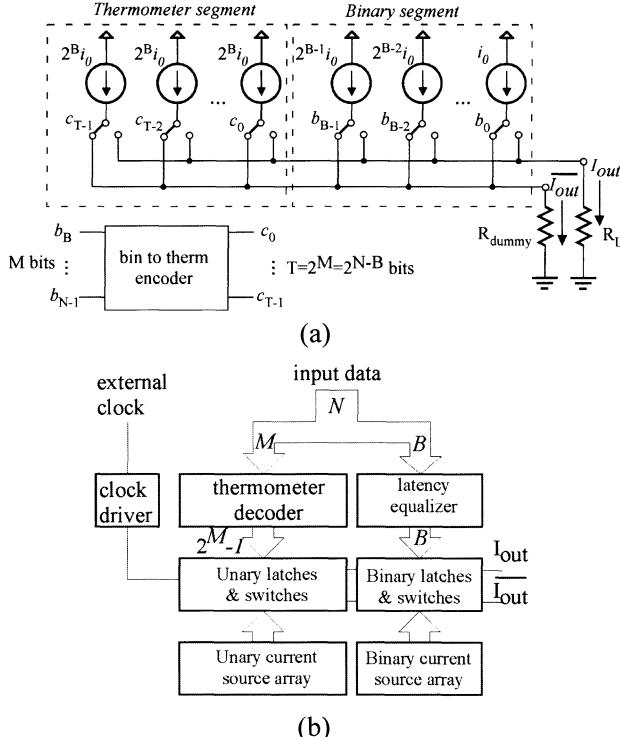


Figure 3.3. (a) Ideal segmented current-steering segmented DAC (b) Architecture-level scheme of current-steering segmented DAC

The performance of the DAC is specified through static parameters (Integral Non-Linearity (INL), Differential Non-Linearity (DNL) and parametric yield) and dynamic parameters (glitch energy, settling time and SFDR) [7]. The design of a current-steering DAC requires as a key architectural decision the selection of the optimum segmentation ratio (M over N). Table 4.1 lists those performance parameters and other information for several previously reported DAC designs.

For the sake of comparison, Figure 3.4 disposes the previous DAC designs in a resolution vs update frequency representation, including information about the implementation technology. The requirements of different current applications appear superimposed.

Table 4.1 State of the art comparison of reported current-steering segmented DAC performance

	Ref	Res. bits	Seg.	INL LSB	DNL LSB	Samp Freq MHz	THD dB	SFDR dBc	Input Freq. MHz	Glitch Ener.	Power mW	Area mm ²	Year	Process
1	[8]	8	6+2	0.3	0.1	150	NA	NA		30pVs	180	4.26	1990	CMOS 1.4μm
2	[27]	8	6+2	<0.5	NA	130	-41	NA	13.5	50pVs	150	0.5	1991	CMOS 1.0μm
3	[42]	8	5+3	<0.5	<0.3	350	NA	NA		30pVs	170	1.6	1995	CMOS 0.5μm
4	[9]	8	3+5	0.125	NA	1000	NA	NA		2pVs	1250	NA	1998	Bipolar
5	[29]	10	7+3	0.5	NA	70	NA	NA		60pVs	170	3.77	1991	CMOS 1.0μm
6	[10]	10	5+5	0.35	0.43	250	NA	60	125	NA	17	0.2	1998	CMOS 0.5μm
7	[11]	10	5+5	0.43	0.7	325	NA	NA		<30pV s	NA	NA	1998	CMOS 0.7μm
8	[12]	10	7+3	0.79	0.69	70	-49	55	1	NA	120	6.20	1999	CMOS 0.8μm
9	[13]	10	8+2	0.43	0.17	100	NA	52	10	NA	NA	3.60	2000	CMOS 0.35μm
10	[14]	10	8+2	0.24	0.09	250	NA	68 (57)	20 (120)	NA	100	1	1998	CMOS 0.5μm
11	[32]	10	8+2	0.2	0.1	500	NA	51	240	NA	125.4	0.6	1998	CMOS 0.35μm
12	[33]	10	6+4	<0.2	<0.15	1000	NA	61.2	490	NA	110	0.35	2001	CMOS 0.35μm
13	[15]	12	8+4	0.8	0.2	65	NA	NA		50pVs	71.7	5	1998	CMOS 0.6μm
14	[16, 22]	12	6+2+ 4	0.6	0.3	300	NA	NA		1.9pV s	NA	1.92	1998	CMOS 0.5μm
15	[17]	12	8+4	<0.5	NA	200	NA	59	2	0.8pV s	140	14	1998	CMOS 0.5μm
16	[34]	12	7+5	0.3	0.25	500	NA	62	125	NA	63	1	2001	CMOS 0.35μm
17	[18]	14	4+10	1	0.9	100	NA	87	2.03	0.5pV s	650	16.73	1997	BiCMOS
18	[2]	14	4+5+ 5	<0.5	<0.5	100	NA	74	8.5	NA	750	14.43	1999	CMOS 0.8μm
19	[30]	14	8+6	<0.3	<0.2	150	NA	61	5	NA	300	13.12	1999	CMOS 0.5μm
20	[19, 35]	14	4+5+ 5	0.32	0.25	200	NA	50 (71)	90 (17)	NA	210	11.83	2000	CMOS 0.35μm

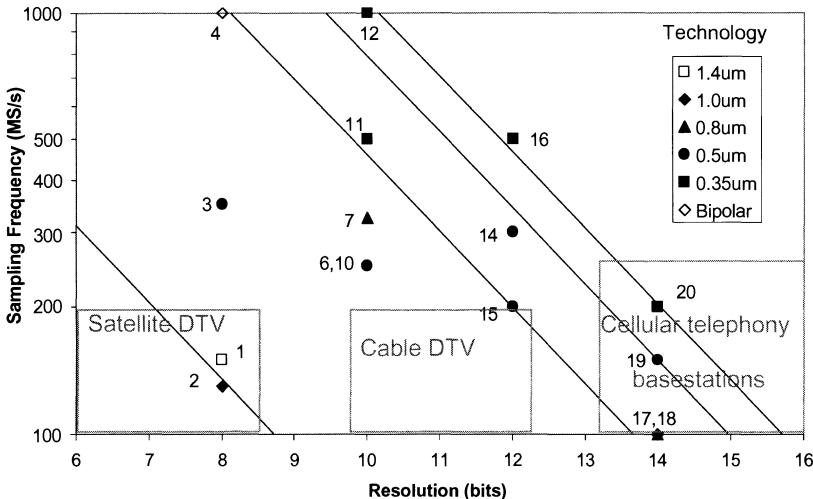


Figure 3.4. Comparison of sampling rate *versus* resolution behavior of recently reported current-steering segmented DACs

3.2 ERROR SOURCES IN CURRENT-STEERING D/A CONVERTERS

Starting from the ideal structure of a current-steering segmented DAC –Figure 3.3(a)-, in the following section it is presented an analysis of the most relevant basic errors affecting this structure, namely the random mismatch among current cells and their non-zero output conductance, as well as their impact over both static and spectral performance metrics, which are previously reviewed.

3.2.1 Static and spectral performance metrics

Static performance is described in terms of indexes representing deviations from the ideal digital input-analog output transfer characteristics. When the steps of the transfer characteristic are unequal, the converter is affected of both DNL and INL, classical static indexes which are indicative of distortion.

Differential Nonlinearity is defined as the deviation of the difference between analog outputs corresponding to adjacent codes referred to its ideal size (1 LSB), and may be expressed normalized in terms of LSBs or not.

$$DNL_k^{\text{norm}} = \frac{s_{k+1}^{\text{real}} - s_k^{\text{real}} - \Delta}{\Delta}, \quad DNL_k = s_{k+1}^{\text{real}} - s_k^{\text{real}} - \Delta. \quad (3.4)$$

Integral Nonlinearity is defined as the difference between the actual analog output and its ideal value, hence being described as the accumulation of previous Differential Nonlinearity errors [3,4].

$$INL_k^{norm} = \frac{s_k^{real} - s_k^{ideal}}{\Delta} = \sum_{j=1}^k DNL_j^{norm}. \quad INL_k = s_k^{real} - s_k^{ideal} = \sum_{j=1}^k DNL_j \quad (3.5)$$

To assure monotonicity, conditions $DNL \leq 0.5$ LSB and $INL \leq 1$ LSB have to be fulfilled.

3.2.2 Mismatch in the current cell

In order to derive the statistical model for the design of the basic current cells, consider that the current provided by each cell i_k is liable of being modeled by a Gaussian random variable described by:

$$E[i_k] = i_0 = \frac{I_{FS}}{2^N} = 1 \text{ LSB}, \quad \sigma_{i_k} = \sigma_{i_0}. \quad (3.6)$$

Recalling that the integral nonlinearity parameter (INL) is defined, for the k -th code and without normalization, as the error between the real current and the ideal current:

$$INL_k = I_k - I_k^{ideal} = I_k - \sum_{j=1}^k i_j = I_k - k i_0. \quad (3.7)$$

The worst case for the INL, independently of whether the architecture consists of binary-weighted or unary-weighted sources (thermometer coded), corresponds to the midcode transition:

$$INL_{max} = INL_{2^{N-1}} = I_{2^{N-1}} - 2^{N-1} i_0. \quad (3.8)$$

Under the hypothesis that each basic source is liable of being modeled as an independent process, the real current is to be modeled by means of a random process described by the following mean and variance:

$$I_{2^{N-1}} \longrightarrow \begin{cases} E[I_{2^{N-1}}] = 2^{N-1} i_0 \\ \sigma_{I_{2^{N-1}}} = \sqrt{\sum_{j=1}^{2^{N-1}} \sigma_{i_j}^2} = \sqrt{\sum_{j=1}^{2^{N-1}} \sigma_{i_0}^2} = \sqrt{2^{N-1} \sigma_{i_0}^2} = \sqrt{2^{N-1}} \sigma_{i_0} \end{cases}, \quad (3.9)$$

resulting that the worst INL –midcode- is modeled by a random variable centered in the origin, described by the following expression and depicted in Figure 3.5 (a).

$$INL^{max} \longrightarrow \begin{cases} E[INL^{max}] = 0 \\ \sigma_{INL^{max}} = \sqrt{2^{N-1}} \sigma_{i_0} \end{cases}. \quad (3.10)$$

If the probability of proper operation of the converter (under criterion error < 0.5 LSB) to be better than 99% (yield) is to be guaranteed, the random variable must verify:

$$p\left(INL^{max} \leq \frac{1}{2} LSB\right) = p\left(INL^{max} \leq \frac{1}{2} i_0\right) = 99\%, \quad (3.11)$$

and considering that for the gaussian function the following condition holds:

$$p(X \leq \alpha\sigma) = 99\% \Leftrightarrow \alpha \approx 2.5, \quad (3.12)$$

results in:

$$2.5\sqrt{2^{N-1}} \sigma_{i_0} = \frac{i_0}{2}, \quad (3.13)$$

finally yielding a relation between the relative precision of each current cell and the resolution of the DAC in number of bits.

$$\frac{\sigma_{i_0}}{i_0} = \frac{1}{\sqrt{2^{N+1}}} \cdot \frac{1}{2.5}. \quad (3.14)$$

Figure 3.5 (b) represents the parametric yield as a function of the LSB source precision for INL < 0.5 LSB. It can be observed in the figure that for a 90% yield the required precision for the current cell (which is ultimately related to its area, as it will be discussed in the subsequent section) for 14, 12, 10 and 8 bits is, respectively, 0.475 %, 0.95 %, 1.9 % and 3.8 %.

Although the previous analysis has been restricted to a design based on the INL specification (which is independent of segmentation), it is immediate to derive the DNL as a function of mismatch in the current cell, which is a much less stringent specification to fulfill. For the thermometer-code case, it is simply due to the connection of a single source, so that it is given by its relative precision:

$$DNL^{max} = \frac{\sigma_{\Delta I}}{i_o} = \frac{\sigma_{i_o}}{i_o} \text{ in } LSB \text{ units.} \quad (3.15)$$

For segmented DAC converters the maximum DNL is limited by the most significant bit B :

$$DNL^{max} = \frac{\sigma_{\Delta I}}{i_o} = \sqrt{2^B - 1} \frac{\sigma_{i_o}}{i_o} \text{ in } LSB \text{ units.} \quad (3.16)$$

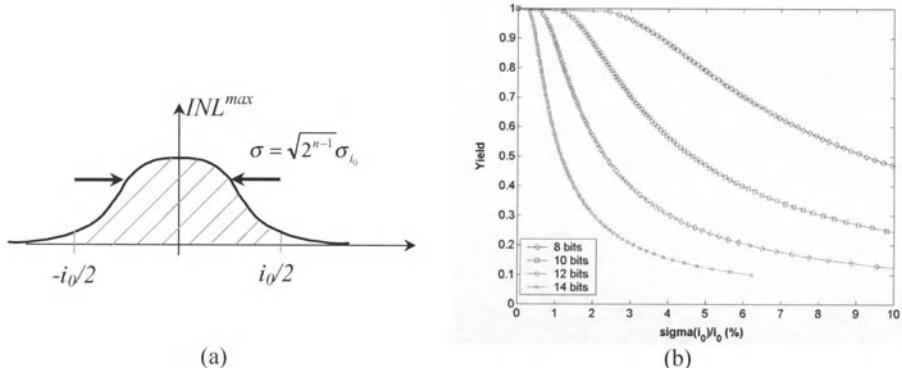


Figure 3.5. (a) Random distribution of the maximum (midcode) INL parameter
(b) Yield as a function of relative precision of the basic current source

3.2.3 Finite output resistance

Considering that the DAC architecture model consists of either a set of $2^N - 1$ LSB current cells operating in parallel or a binary-weighted version, since either model provides the same result, in the following, the nonlinear impact of the non-ideal finite output resistance of each current cell is examined. As shown in Figure 3.6, the equivalent model of the DAC explicitly shows how the inclusion of output resistance for each individual cell results in a code-dependent finite output resistance for the whole DAC, which reverts in a nonlinear effect.

From the equivalent modeling depicted in Figure 3.6, the output current flowing through the load is given by

$$i_L = k i_o \frac{1}{1 + k \left(\frac{R_L}{R_o} \right)}, \quad (3.17)$$

where the denominator clearly indicates the nonlinear effect as a consequence of the output resistance. To quantify this nonlinearity, the direct application of the definition of DNL in eq. (3.4) and taking into account eq. (3.17) yields

$$DNL_k = i_k^{real} - i_{k-1}^{real} - i_o = i_o \left[\frac{1}{1 + k \frac{R_L}{R_o}} - \frac{1}{1 + (k-1) \frac{R_L}{R_o}} \right] - i_o \equiv \frac{i_o}{(1 + k \frac{R_L}{R_o})^2} - i_o \quad (3.18)$$

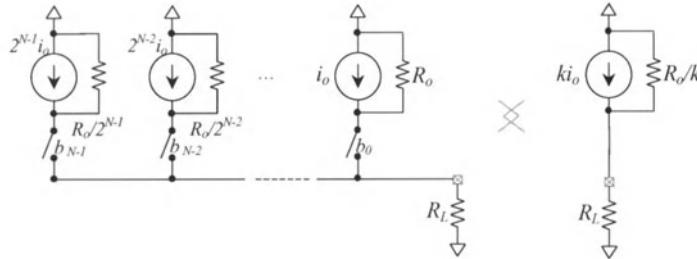


Figure 3.6. Ideal model of a current-scaling DAC taking into account output resistances

Complementary, the integral nonlinearity is obtained directly from eq. (3.18) after aggregation of DNL terms according to (3.5). The dependence of both static performance indexes as a function of the resistor ratio R_o/R_L is represented in Figure 3.7. The degradation of static performance when the output resistance is reduced can be observed.

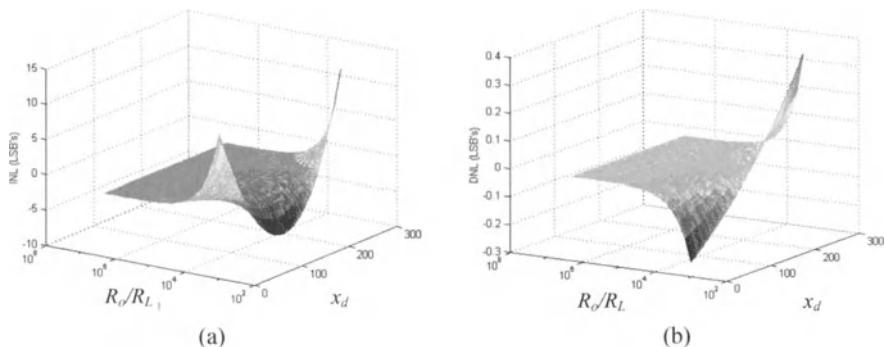


Figure 3.7. Static performance of an example 8 bit converter under nonideal output resistance (a) INL (b) DNL

Although it is inferred from the previous Figure 3.7(b) that for a reasonably high output resistance monotonicity is guaranteed, spectral performance of the DAC converter may be compromised, and this deserves further analysis. Subsequently, spectral performance is analyzed through the SFDR measure, which is one of the most used spectral indexes, being defined as the ratio of the power of the signal to the power of the largest spurious within the frequency band of interest, being thus expressed as

$$SFDR(dBc) = 10 \log \left(\frac{X_{fund}^2}{X_{spur}^2} \right). \quad (3.19)$$

By considering that the input digital code stream corresponds a sinusoidal signal:

$$k(t) = K_{dc} + K_{ac} \sin(\omega t + \theta), \quad (3.20)$$

substituting the previous expression in eq. (3.17), and after retaining only the *ac* component and representing its Taylor series expansion, a quotient of the fundamental harmonic term and the second harmonic is obtained as:

$$SFDR = \left[\frac{1 + \frac{R_L}{R_o} K_{dc}}{\frac{R_L}{R_o} K_{ac}} + \sqrt{\left(\frac{1 + \frac{R_L}{R_o} K_{dc}}{\frac{R_L}{R_o} K_{ac}} \right)^2 - 1} \right]^2 \xrightarrow{\frac{R_L \downarrow \downarrow}{R_o}} 2 \left(\frac{\left(\frac{R_L}{R_o} \right)^{-1} + K_{dc}}{K_{ac}} \right)^2. \quad (3.21)$$

The SFDR is much more sensitive to K_{ac} than K_{dc} . Provided that a full-scale sinusoid is applied $K_{ac}=K_{dc}=2^{N-1}$ and $R_o/R_L >> 2^N$, the previous expression can be simplified to:

$$SFDR(dBc) \cong 20 \log \left(\frac{R_o}{R_L} \right) - 6(n-2). \quad (3.22)$$

Expression (3.21) is plotted in Figure 3.8, where it can be seen that doubling the resistance ratio reduces SFDR in 6dB.

It is worth indicating that fully-differential signaling results in a cancellation of even harmonics, so that resistance requirements are relaxed, the SFDR is clearly improved, and can be modeled after an analogous derivation as:

$$SFDR(dBc) \cong 40 \log\left(\frac{R_o}{R_L}\right) - 12(n-2). \quad (3.23)$$

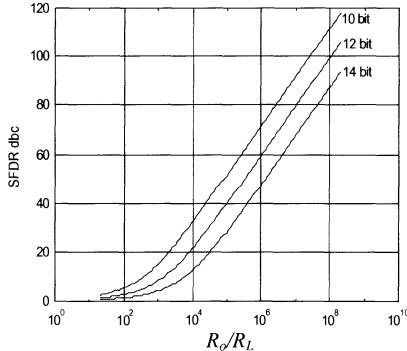


Figure 3.8. Spectral performance through SFDR as a function of R_o/R_L

3.3 CIRCUITAL IMPLEMENTATION OF CURRENT-STEERING D/A CONVERTERS

The transistor-level circuit implementation of the previously presented architectures for current-steering DAC converters is presented in this section. In particular, two possible topologies for the basic cells are presented and briefly analyzed, more specifically the relation of their MOSFET design parameters (width, length and biasing) to both mismatch and output impedance of the current cell. Subsequently, dynamic errors and their ultimate impact on spectral performance are presented.

3.3.1 The current cell

3.3.1.1 Single transistor current cell

The simplest topology implementing the current cell consists of a single MOS transistor biased with constant gate-source voltage and operating in saturation. The current-steering switch is implemented straightforwardly by means of MOS transistors operated as switches with complementary activation gate signals. The resulting topology is represented in Figure 3.9, where the different elements are denoted as the current source (CS) transistor and the two complementary switch transistors (SW, \overline{SW}). This topology can alternatively be interpreted as a

differential pair which steers the biasing current as a response to input on-off stimuli.

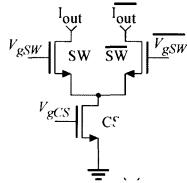


Figure 3.9. Single transistor current cell circuit

3.3.1.2 Mismatch and output impedance in the single transistor current cell

The previous expression obtained for the required relative precision of the current cell eq. (3.14) is directly related to the design parameters and spatial distribution of transistor CS, which impact on mismatch among different current cells. The stochastic model is described in the following. Recall that any parameter P of a transistor subject to statistical variations can be modeled by:

$$\sigma_P^2 = \frac{A_P^2}{WL} + S_P^2 D^2, \quad (3.24)$$

where the first term corresponds to fast local gradients (or random errors), while the second term corresponds to slow global gradients (or systematic errors) [20]. As a first approximation, it will be considered that the second term is negligible under the assumption that transistors implementing the current sources are sufficiently close. Nevertheless, for high resolutions, the transistor array becomes large and this statement does not hold, hence requiring specific compensation techniques at an architecture and layout level, which are the purpose of section 4.4. The first term is solely determined by technology and transistor sizing. In particular, a commonly accepted model reveals a random dependence of both threshold voltage, the gain parameter as well as body effect parameter:

$$\sigma_{V_T} = \frac{A_{V_T}}{\sqrt{WL}} \quad \frac{\sigma_\beta}{\beta} = \frac{A_\beta}{\sqrt{WL}} \quad \frac{\sigma_\gamma}{\gamma} = \frac{A_\gamma}{\sqrt{WL}}. \quad (3.25)$$

These relations can be accounted for to obtain the relative precision of the current supplied by the transistor (taking also into account a simple quadratic

transfer for the transistor in saturation and neglecting body effect as corresponds to transistors implementing current sources):

$$\left. \begin{aligned} \sigma_{V_T} &= \frac{A_{V_T}}{\sqrt{WL}} \\ \frac{\sigma_\beta}{\beta} &= \frac{A_\beta}{\sqrt{WL}} \\ \frac{\sigma_\gamma}{\gamma} &= \frac{A_\gamma}{\sqrt{WL}} \end{aligned} \right\} \xrightarrow{I=\frac{\beta(V_{GS}-V_T)^2}{2\gamma}, \gamma=0} \left(\frac{\sigma_I}{I} \right)^2 = \left(\frac{\sigma_\beta}{\beta} \right)^2 + \frac{4\sigma_{V_T}^2}{(V_{GS}-V_T)^2} = \frac{A_\beta^2}{WL} + \frac{4A_{V_T}^2}{WL(V_{GS}-V_T)^2}. \quad (3.26)$$

From the previous expression, the minimum area guaranteeing the required relative precision is derived:

$$WL = \frac{A_\beta^2 + \frac{4A_{V_T}^2}{(V_{GS}-V_T)^2}}{\left(\frac{\sigma_I}{I} \right)^2}. \quad (3.27)$$

This expression provides the area of the current cell, but a degree of freedom is retained. Note that the derivation is independent of the absolute value of the current flowing through current sources, which is, in turn, related to the speed and output dynamic range of the cell. Hence, the aspect ratio is determined by the output voltage margin required for the transistor to operate in saturation (considering ideal switches):

$$V_{GS} - V_T = \sqrt{\frac{i_0}{K' \frac{W}{L}}} = V_{DD} - \Delta V_{out}^{max}. \quad (3.28)$$

By considering the relation between the basic cell current level and the full scale current margin, and that this full scale is related to voltage full scale range via de output resistance R_L :

$$\begin{aligned} i_0 &= \frac{I_{FS}}{2^N}, \\ I_{FS} &= \Delta V_{out}^{max} R_L \end{aligned} \quad , \quad (3.29)$$

the aspect ratio is finally obtained:

$$\frac{W}{L} = \frac{1}{R_L 2^N K'} \frac{\Delta V_{out}^{max}}{(V_{DD} - \Delta V_{out}^{max})^2}. \quad (3.30)$$

Once the current cell is designed to fulfill the mismatch requirement, the other factor affecting the distortion is the output resistance, as studied in previous subsection 4.2.3. The output resistance of the current cell transistor, which is roughly proportional to its channel length, is determined by the precedent mismatch procedure, and it is not sufficient to fulfill requirements even for low resolutions. One circuit option retaining the simplest circuit topology is to operate the switch transistors in their saturation region to provide cascode effect when in ON state so as to improve output resistance:

$$r_o = r_o^{CS} \left(g_m^{SW} r_o^{SW} \right). \quad (3.31)$$

3.3.1.3 Cascoded current cell

The maximum achievable output resistance of the simple topology, even with cascode operated switches (3.31) does not suffice in high-resolution applications, which require to resort to the enhanced topology shown in Figure 3.10. The circuit includes an additional cascode transistor (CAS) that increases the output impedance to fulfil the SFDR specification for high resolutions.

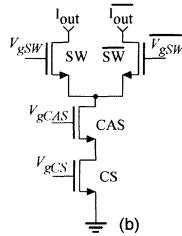


Figure 3.10. Cascoded current cell circuit

The circuit in Figure 3.10 exhibits the output resistance of a double-cascode structure when switches are operated in the limit of saturation.

$$r_o = r_o^{CS} \left(g_m^{CAS} r_o^{CAS} \right) \left(g_m^{SW} r_o^{SW} \right). \quad (3.32)$$

If properly designed, this structure is capable of providing output resistances in the $\text{G}\Omega$ range, thus suiting low-distortion applications.

3.3.1.4 Frequency-dependent current cell output impedance

Previous expressions for output resistances of current cell implementations model only static behavior. However, as it is shortly discussed in the following,

the dynamic frequency-dependent impedance of the cell has to be taken into account to avoid severe reduction in SFDR.

When considering for the basic current cell in Figure 3.9 an output capacitance C_o from the drain of the CS transistor to ground, the output impedance of this current cell presents a pole followed by a zero, hence reducing the impedance when increasing frequency:

$$Z_o(j\omega) = r_o^{CS} \left(g_m^{SW} r_o^{SW} \left(\frac{1 + \frac{j\omega C_o}{g_m^{SW}}}{1 + j\omega C_o r_o^{CS}} \right) \right). \quad (3.33)$$

Given the fact that capacitance C_o is obtained as a result of the design of the CS transistor and that there are no degrees of freedom, the decrease in output resistance reduces the SFDR for high input frequencies (through equation (3.22)), as illustrated by Figure 3.11, where a 12 bit DAC is simulated with a pole at 1 MHz from $10 \text{ G}\Omega$ at DC. In addition, this simulation assumes a load impedance with a pole at 200 MHz.

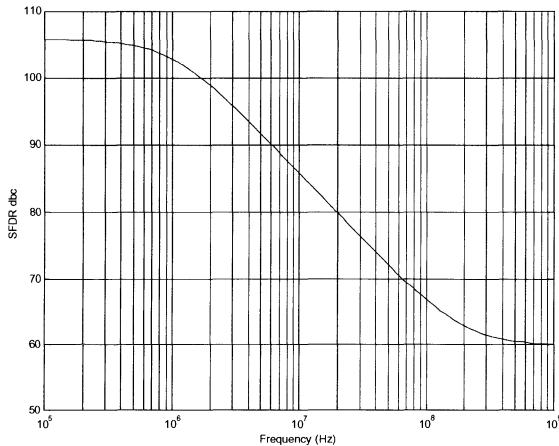


Figure 3.11. Effect of frequency-dependent output impedance on SFDR

Capacitance C_o is mainly driven by layout from routing wires, accounting for the fact that the matrix of the current cells is spatially separated from the matrix of switches so as to reduce slow gradient mismatch. In order to retain reasonable SFDR, a solution has to be obtained at design level, and this is provided by the extra degrees of freedom of the cascaded current cell in eq. (3.10), whose

frequency-dependent impedance can be properly designed to attain higher values at higher frequencies, as discussed in [21].

3.3.2 Dynamic errors

Despite previous analysis and modeling mainly consider static or low-frequency operation of the converter, when high frequency operation close to the full Nyquist range is considered, several dynamic errors that influence performance appear. Not only these errors are frequency dependent but they are signal-dependent as well, hence reducing spectral performance.

3.3.2.1 *Switch transistors behavior*

The switch activation should be ideally complementary to completely steer the current and to assure that the current cell finds a low-ohmic path at any time instant, or either the voltage at the internal node raises provoking important delays and glitches. At circuit and layout level, this requires a synchronization block just before the switches.

The cascode topology reduces the clock feedthrough from the switches to the drain of the CS thus reducing the glitch energy. A driver circuit with a reduced swing placed between the latch and the switch reduces the clock feedthrough to the output node as well [22], as will be discussed in more detail in section 4.6.

3.3.2.2 *Glitches*

Glitches at the output of a DAC are peak deviations during output steps due to mismatching in the switching times of different bits in a binary weighted architecture, which result in the momentary application of erroneous codes. Glitches are difficult to model and are usually characterized by their area, as a measure index of amplitude and time duration.

The glitch energy is determined by the number of binary bits B , being the optimum architecture in this sense a totally unary DAC. However this is unfeasible in practice due to the large area and delay that the thermometer decoder would exhibit. The minimization of the glitch energy is then bypassed to the circuit level design of the switch & latch array and current source cell.

3.3.2.3 *Settling time*

Since the output of a DAC converter is an analog signal, when it is operated close to its speed limit, non-negligible settling time has to be accounted for. Settling time is defined as the time taken by the circuit structure to reach a certain percent of the ideal final value, hence determining the highest speed of the circuit.

Settling time is usually described by analyzing the linear circuit resulting from the inclusion of parasitic capacitances. A dominant pole approximation (with associated time constant τ) can predict the maximum operating frequency f_s of the DAC as a function of the resolution in bits N :

$$\tau < \frac{1}{f_s(N+1)\ln 2}. \quad (3.34)$$

However this result is optimistic since actual nonlinear settling due to code dependent capacitances results in distortion.

3.3.2.4 Clock jitter and dI/dt noise

When multiple bits of the thermometer code segment (TCS bits) of the input signal switch simultaneously, the decoder will generate dI/dt noise (or switching noise) at the internal power supply nodes. The amount of dI/dt noise roughly depends on the effective power and ground pins package inductance, the on-chip built-in decoupling capacitance and the digital current pulse shape. The later is determined by the logic signal rise times, the saturation current of the gates output transistors and the number of simultaneously switching gates [23]. In this sense, there is a relation between the TCS bits simultaneously switching and the amount of switching noise that is generated. The maximum noise will be produced when all the TCS bits make a simultaneous transition. Thus, a larger M (the number of input bits assigned to the TCS) increases the maximum dI/dt noise that is generated by the decoder in the worst case. For the sake of simplicity it can be assumed that the amplitude of the dI/dt noise spike increases linearly with the number of simultaneously changing bits. This is justified by the fact that the circuits with highest fan-out are the inverters that complement the input bits and distribute them to several decoder gates. For example, in the vicinity of the crossover point of a sinusoidal input, one sample should be 0111111111111111 and the next should be 1000000000000000. If the thermometer code segment has a length $M = 8$ bits, there will be a total of 8 simultaneous switching large fan-out inverters at the input of the decoder. Meanwhile, at the decoder output only one of the 256 bits will turn on.

dI/dt noise is responsible of provoking jitter. Note that the exact time at which the analog output signal switches is fixed by the clock signal edges. The clock driver generates the internal clock signal from an external clock. The digital input codes are supplied synchronized with the external clock. The input bits changes generate dI/dt noise at the internal power supply nodes, as explained before. That noise is either coupled directly to the clock driver if it shares the same digital power supply or it is coupled through the substrate if it does not. The dI/dt and/or substrate noise affects the delay of the output driver, generating clock-jitter [24].

Figure 3.12 shows the simulated delay change of a clock driver due to dI/dt noise generated by the decoder when multiple bits of the input code simultaneously change, for a particular $0.5\mu\text{m}$ technology. The simulation considers an on-chip decoupling capacitance and resistance of 5 pF and 1Ω , respectively, and the power supply package pins inductance is varied from 3 nH to 10 nH . The triangular current pulse used to model the decoder activity is 500 ps wide and its amplitude is varied from 5 mA to 100 mA . The input data is changed 3 ns before the clock rising edge. Even if the current pulse due to the input data change is not simultaneous with the external clock edge (as is the case here), the damped oscillation generated by the dI/dt noise pulse in the package-chip RLC resonant circuit [25] affects the internal clock edge delay. Different noise levels from sample to sample produce different delay changes in the clock driver and this is the cause of the internal clock-jitter. This effect is much more pronounced than that due to thermal noise, as shown in [24].

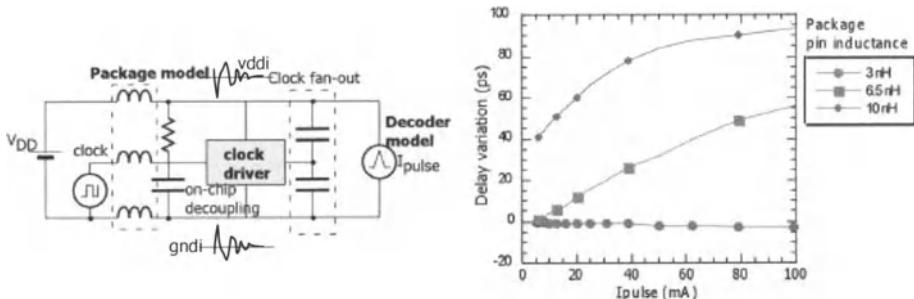


Figure 3.12. Clock signal delay changes caused by dI/dt noise.

In the analysis of a subsequent section it will be used a linear expression relating the clock-jitter amplitude and the number of input bits simultaneously switching based on the results of Figure 3.12. Only the changes in the TCS bits are relevant. In this way, an increase in the size of the thermometer segment M will raise the maximum clock-jitter and produce a lower SFDR, as shown below.

3.3.2.5 Spectral performance vs dynamic errors

The previous dynamic errors result in nonlinearities that severely affect the frequency-domain performance of high-speed DAC converters, hence imposing a limit in the maximum operable frequency for a given performance. The modeling of their impact over SFDR and other indexes is complex and is a subject of current research, and thus they are usually simulated or measured. In following sections, some design hints pursuing the reduction of these effects are discussed.

3.4 ARCHITECTURES FOR SYSTEMATIC ERROR COMPENSATION

Process-induced variations are an important consideration in the design of integrated circuits and moreover in high-speed D/A converters where the accuracy can be severely limited by those variations. In addition to the fine-grain random parameter fluctuations considered in previous sections there are also errors developed at chip level that show a slow gradient profile, which corresponds to the distance-dependent term in [4.24]. They came both from the wafer level gradients developed during the process, originating in non-uniformities of the resist thickness or stepper-lens aberrations [26], or from the chip level where stress and temperature gradients are developed. Other sources of systematic errors at chip level are edge effects originating in the different geometrical environment found in some parts of the layout, for example, around the edges of the current source array [27].

These process-induced variations produce systematic parameter fluctuations across the surface of the chip. The impact of this systematic parameter fluctuations is more severe in regular structures of theoretically equal devices placed in array structures, as is the case of the current cells array of the current-steering D/A converters. In that case, the systematic parameter fluctuations give rise to a mismatching between the current generated by two transistors separated by a distance D , which is proportional to this distance D [20]. In order to minimize the error in the output transfer function of the DAC some techniques can be used to compensate for the systematic parameter fluctuations.

3.4.1 One dimensional switching sequences for systematic error compensation

The techniques presented in this subsection are suitable for row-column decoding architectures. The simplest model for the systematic errors in the current cell array is a linear one-dimensional gradient in the horizontal direction of the array and another one in the vertical direction. The two one-dimensional gradients are considered to be independent. This situation is represented in Figure 3.13, where a 8-bit DAC segmented into six thermometer bits and two binary bits is shown. When the unary current sources are switched in an ordered way the error accumulates along the x axis producing an INL that increases with the input code. In order to compensate for the effects of the linear gradient error, other switching sequences can be used instead. Figure 3.14 shows the same 8-bit D/A converter where the rows and columns are switched using a *symmetrical switching sequence*, which compensates 1D errors of symmetrically placed current sources. This is done independently along the two dimensions of the array [28].

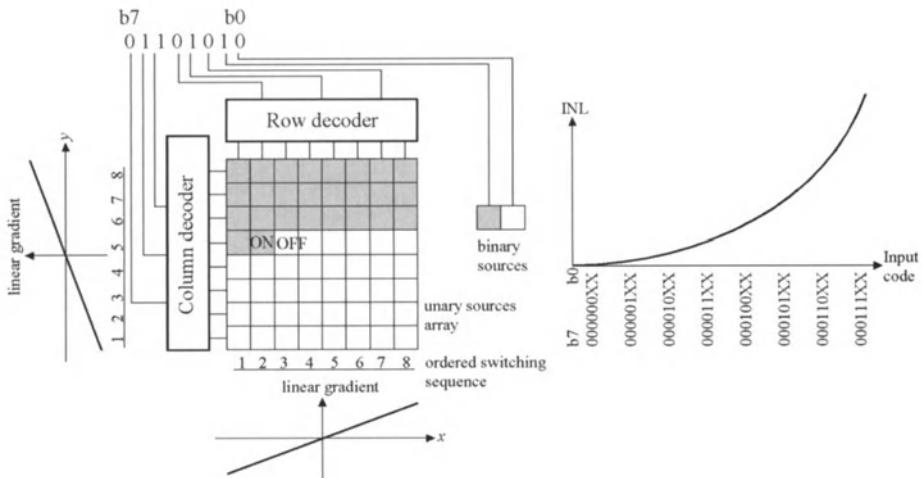


Figure 3.13. Ordered switching sequence for a 8-bit current-steering segmented DAC.

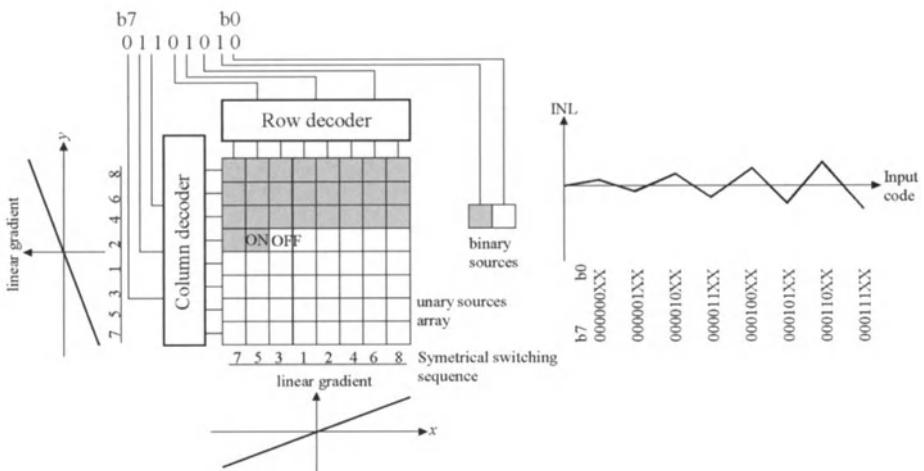


Figure 3.14. Symmetrical switching sequence to compensate for linear gradient errors in the current array

The linear gradient error is only a simple approximation of the real gradient errors. A more accurate yet also simplified model includes two one-dimensional components of the gradient error along the two directions of the current source array. The first component is a linear gradient term and the second is a symmetrical component centered in the middle of the row (or column) of the

array. In order to compensate for the two gradient error terms a *hierarchical symmetrical switching sequence* can be used, as shown in Figure 3.15 for a 10 bits segmented DAC with 7 thermometric (implemented in a 8×15 unary source array) and 3 binary bits [29].

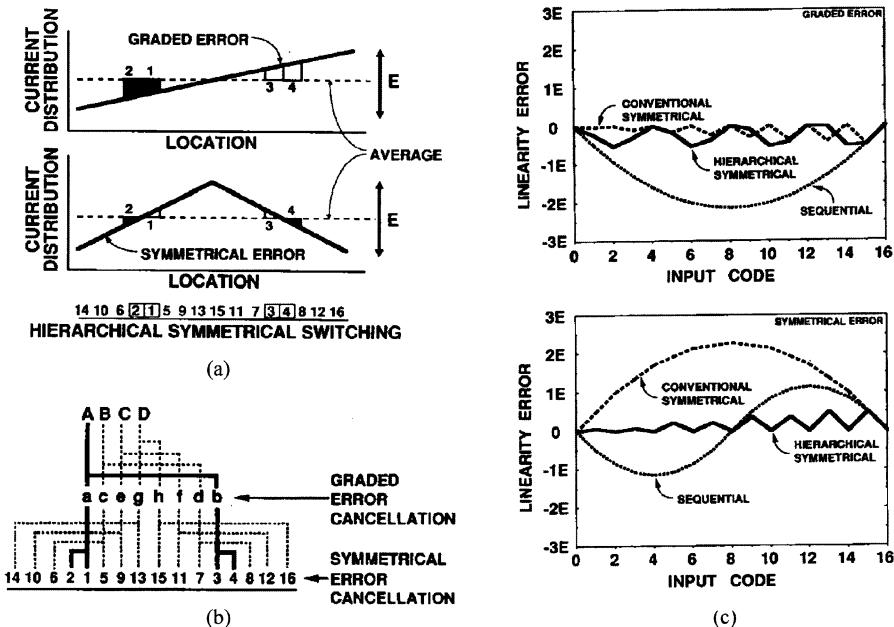


Figure 3.15. (a) Linear (graded) and symmetrical one-dimensional gradient error profiles. (b) Hierarchical symmetrical switching sequence to compensate for both one-dimensional linear and symmetrical gradient errors. (c) Comparison of linearity errors (INL) obtained by the sequential (ordered), symmetrical and hierarchical symmetrical switching sequences in the presence of one-dimensional linear and symmetrical gradient errors. Figures from [29], reprinted with permission.

3.4.2 Two dimensional switching sequences for systematic error compensation

Considering systematic gradient errors to be independent, one dimensional error profiles in the two dimensions of the current source array is a simplification that may work for medium resolution (up to 10 bits). For larger resolutions, however, this simplification - and the systematic error compensation techniques presented in the previous subsection based on it - leads to unacceptable linearity

errors. Systematic gradients have actually a more complicated shape. They can be approximated by its Taylor series expansion around the center of the current source array (the constant term of the series expansion will contribute to an offset or gain error in the output transfer function that has no relevant impact on the DAC static performance). Two-dimensional (2D) switching sequences can be found to compensate for the first two terms of the error series expansion (the linear and the quadratic term), as shown in Figure 3.16. The basic idea behind this 2D switching sequence is the same than explained in the previous subsection. The placement of current sources across the current source array is organized in such a way that when they are switched following an increasing input code order, their locations contribute with errors that accumulate and compensate with previously switched-on current sources. The 2D switching sequence is also hierarchical, compensating in one level for the linear term of the gradient error and in the second level for the quadratic term. This approach compensates better for systematic gradient errors than the previous ones but it cannot be implemented using a row and column decoding structure. A single binary to thermometric decoder has to be used for the whole current source array because the switching order sequence is done in a two-dimensional way that is not possible to organize in rows and columns.

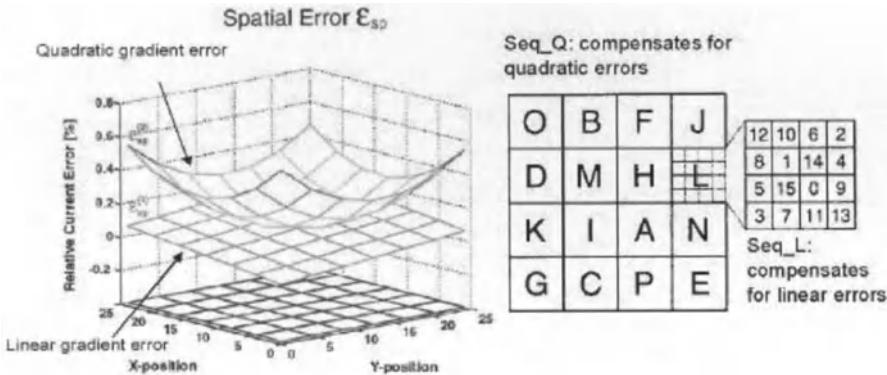


Figure 3.16. 2D Linear and quadratic terms of the gradient error and hierarchical sequence to compensate them in a 14 bits segmented DAC with 8 thermometric (organized in a 16×16 unary current source array) and 6 binary bits. Figures from [30] reprinted with permission.

The 1D and 2D switching sequences shown previously were heuristically derived. Indeed, it is possible to find other switching sequences with similar or better INL minimization performance for the same cases. In [31] a process to find both 1D and 2D optimum switching sequences is presented. The only information required by the algorithms used to find such sequences is the angle

between the linear gradient and the two axis of the current source array (θ), and the relative amplitude of the quadratic term respect the linear term of the gradient error. Actually the sequences presented in [31] are optimum for any θ .

One important consideration to be taken into account when floorplaning the DAC current unary current source array is the size of the basic current cell. The current source transistor and the other basic current cell components (switches, cascode transistor, latch, etc) can be placed all together in each cell of the array or in separate arrays. The first option has the advantage of minimizing the interconnection length between the current source transistor and the switches but then the total current source array area becomes very large, making systematic gradient errors more important. The preferred approach for high-accuracy DACs is to place in the unary current source array only the current source transistors. All the other circuitry is placed in a separate array, as will be explained later in section 3.6.

3.4.3 Other systematic error compensation techniques

In addition to the switching sequence ordering other techniques can be used to further reduce systematic errors.

3.4.3.1 Current-source transistor common-centroid distribution

The current source transistor can be split up in several equal parts (sub-units) and those sub-units distributed in a common-centroid way. This process can be extended recursively yielding to double or triple common centroid distributions of the individual sub-units of each one of the current source transistors. Then the sub-units corresponding to the different current sources are placed using a switching sequence as explained in the previous sub-section. Figure 3.17 illustrates some common centroid distributions with and without switching sequence ordering. As shown in the figure, the binary source transistors are placed in the empty spaces of the unary current source array. They are also split up in several sub-units and placed in a common centroid distribution. All these common-centroid distribution compensate in a two-dimensional fashion for linear and other even order terms of the systematic gradient error. Remaining linear and quadratic terms can be compensated with a 2D switching sequence, as shown in Figure 3.17-(a).

3.4.3.2 Split up common-centroid biasing scheme

This technique has been used in [32] to improve DNL and INL of a 10 bits DAC (8 thermometric and 2 binary bits). In that work the basic current cell includes the current transistor and also the switch, a cascode transistor, the latch and decoding logic. The current cells are placed in the array following the same

1D randomized switching sequence for the row and columns. Then, the array is divided in four quadrants and the current cell transistor of the cells of each quadrant is biased independently. The four biasing voltages are generated with four bias generators. Each of these bias generators is split up in two sub-units and the two sets of four sub-units of the four bias generators are distributed in a common-centroid layout, as shown in Figure 3.18.

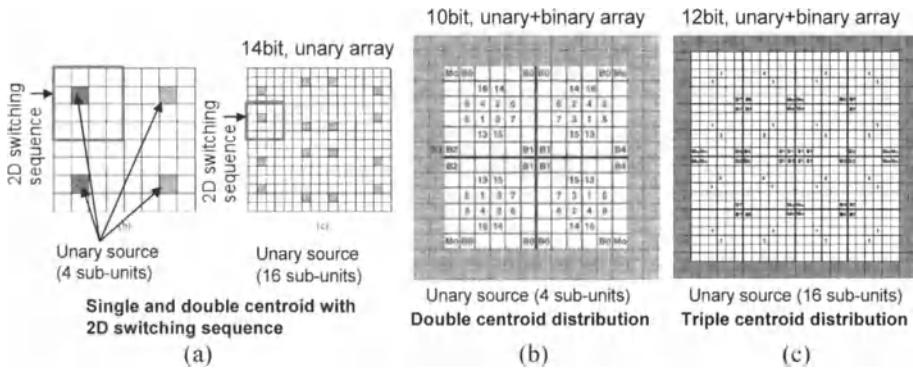


Figure 3.17. Examples of centroid distribution. (a) Single and double centroid distribution in conjunction with a 2D hierarchical switching sequence used in the 14 bits DAC of [30]. (b) Double centroid distribution without switching sequence used in the 10 bit DAC (5 thermometric and 5 binary bits) of [33]. (c) Triple centroid distribution without switching sequence used in the 12 bit DAC (8 thermometric and 6 binary bits) of [34].

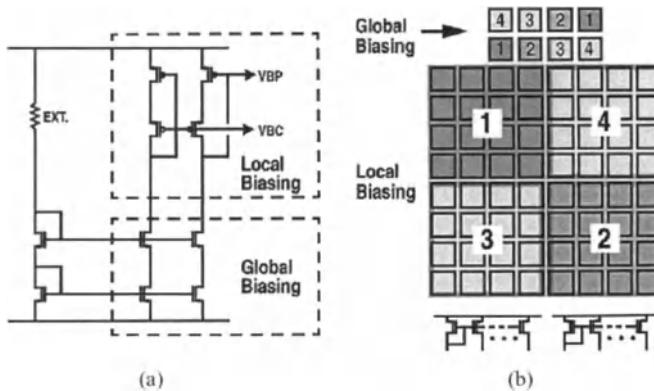


Figure 3.18. (a) Biasing generator circuit and (b) biasing scheme used in [32] to compensate for systematic gradient errors. Reprinted with permission.

3.4.3.3 Self-trimming and dynamic matching

This technique is used to improve the accuracy of high-resolution DACs. It would be very complicated to implement self-trimming for all the bits of the converter. A usual approach [35,36] consists on calibrating only the most significant bits (MSB) of the converter. The required intrinsic accuracy for the lower bits is obtained using one of the previously described techniques for compensating systematic errors. The MSB bits are implemented in a thermometric way. The current sources controlled by the MSB bits have all the same value which is equal to the total value of the less significant section of the converter plus one additional LSB. The MSB section unary current sources are self-trimmed with a feedback loop used to dynamically match them to the lower significant current sources. Each one of the unary current sources of the MSB includes a trimming mechanism that allows adjusting its current value. The control word for the MSB unary current sources trimming is stored in a memory that is updated periodically by the feedback loop.

3.5 OPTIMUM SEGMENTATION IN CURRENT-STEERING SEGMENTED D/A CONVERTERS

In section 4.2 the relation between the static and dynamic performance and the DAC circuit structure and technology parameters has been presented. In the following, this relations will be used to present two distinct design trade-offs to optimize the segmentation of the DAC, *i.e.*, the number of thermometer (M) and binary (B) bits, for a given number of total input bits ($N = M + B$).

3.5.1 Optimum segmentation based on area vs static performance

Firstly, the impact of the degree of segmentation ($M/N \%$) over the static performance and the area of segmented current-steering DAC is analyzed. It has been shown in section 4.3 that the area of a current cell transistor is inversely proportional to its relative accuracy due to random mismatch errors. Systematic mismatch errors are compensated by layout techniques, as explained in section 4. The INL and DNL specification of the converter, which are mainly limited by mismatching, place a constraint in the minimum required accuracy, and correspondingly on the area of the current cell transistors.

The worst-case INL specification of the converter, for a given parametric yield, is related to the accuracy of the LSB current source by:

$$P(INL^{\max} \leq \frac{1}{2}i_0) \Rightarrow \alpha\sqrt{2^{N-1}}\sigma_{i_0} = \frac{i_0}{2}, \quad (3.35)$$

where i_0 is the LSB current, σ_{i_0} its standard deviation, and α is the number of standard deviations of the INL Gaussian distribution required to fulfill the INL specification (less than half an LSB, in this case) in a given percentage of the samples, expressed by the parametric yield.

The worst case DNL in LSB units is found for the mid code transition:

$$DNL^{\max} = \sqrt{2^{B+1} - 1} \frac{\sigma_{i_0}}{i_0}. \quad (3.36)$$

The two previous expressions can be used to find two different requirements for the relative accuracy of the basic current source σ_{i_0}/i_0 , that will determine the minimum area of the current source transistor and hence the total area of the analog section of the converter. The analog area will be ultimately determined by the most restricting of the two specifications, either INL or DNL. The former is independent of the segmentation but the later depends on the number of binary bits. When the number of binary-code bits is increased, the same DNL specification requires a smaller relative accuracy and hence larger current source area.

The total analog area for a N bits segmented current steering DAC can be estimated using the following expressions, for two INL and one DNL specifications:

$$\begin{aligned} INL < 0.5 \text{ LSB} &\Rightarrow \frac{\sigma_{i_0}}{i_0} < \frac{0.5\alpha}{\sqrt{2^{N-1}}} \Rightarrow A_{ana} \propto \frac{1}{\alpha^2 4} 2^{N-1} (2^N - 1) \\ INL < 1 \text{ LSB} &\Rightarrow \frac{\sigma_{i_0}}{i_0} < \frac{\alpha}{\sqrt{2^{N-1}}} \Rightarrow A_{ana} \propto \frac{1}{\alpha^2} 2^{N-1} (2^N - 1), \\ DNL < 0.5 \text{ LSB} &\Rightarrow \frac{\sigma_{i_0}}{i_0} < \frac{0.5}{\sqrt{2^{B+1} - 1}} \Rightarrow A_{ana} \propto \frac{1}{4} (2^{B+1} - 1) (2^N - 1) \end{aligned} \quad (3.37)$$

where A_{ana} is the total analog area of the converter found by multiplying the area of a single LSB current source by the total number of LSB current sources. The constant α is positive. For example, for a parametric yield of 99% α is 2.5. From eq. (3.37) two important conclusions can be drawn:

- Regarding INL area constraint: every extra bit of resolution implies multiplying by four the total analog area.
- Regarding DNL area constraint: an increase in the number of thermometer segment bits (and hence a reduction of the number of binary bits), reduces by a power of two the total analog area.

Increasing the number of thermometer bits M has another impact on the total converter area, however. The number of unary current sources is equal to $2^M - 1$, and every unary current source requires a thermometer decoded control signal, a couple of switches and a latch. The area of the unary current source is accounted in the analog area expressions of eq. (3.37), along with the binary current sources area. However, the additional digital circuitry has also to be considered as it contributes to an important part of the DAC total area. The switch transistors are scaled accordingly with the current sources, but they do not represent a significant part of the total digital area. The most area-consuming block is the latch (including the drivers for the switch transistors). The area of this block is independent on the current source area (either if it stores the control signal for a unary or a binary weighted current source). The latch blocks overall contribution to the DAC digital area is proportional to 2^M . The second important contributor to the total digital area is the binary to thermometer decoder, whose total area is also approximately proportional to 2^M . In conclusion, the total digital area of the converter is proportional to 2^M and therefore, increases exponentially as the number of thermometer segment bits is increased.

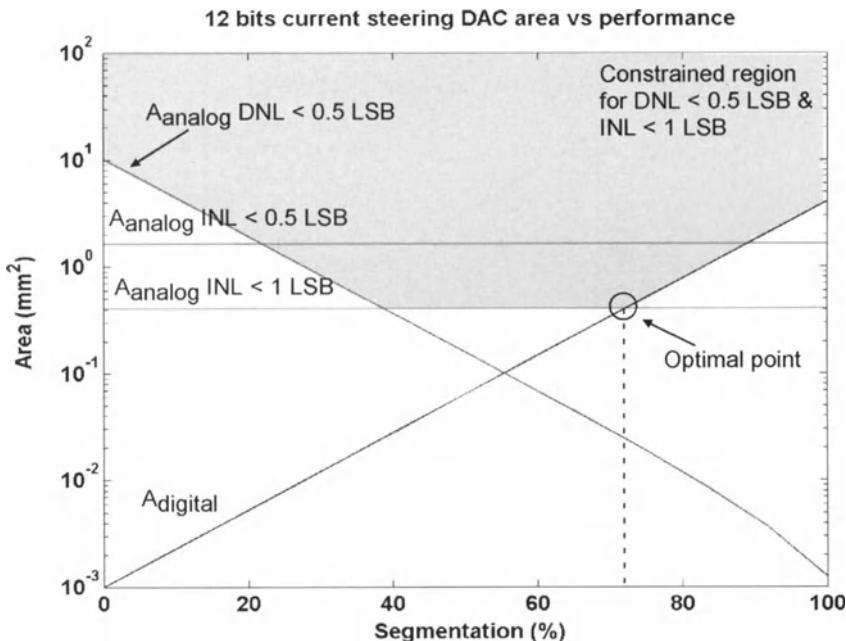


Figure 3.19. Required area versus segmentation ratio to fulfill the two static performance constraints for a 12 bits segmented current-steering DAC.

The total analog and digital area for a 12 bit segmented current-steering DAC is shown in Figure 3.19, considering the two static performance constraints (INL and DNL) and all the possible segmentation ratios. The graphs have been drawn from actual data of a 12 bit converter designed using a $0.35\text{ }\mu\text{m}$ CMOS process and are in good agreement with previously published works [32]. According to these results it is apparent that DNL analog area constraint dominates for small segmentations, not used in practice. Hence the area *versus* static performance trade-off is established between the INL analog area constraint, which is independent of the segmentation, and the digital area that increases exponentially with the segmentation ratio. The optimal point is obtained by taking into account that, as discussed, the glitch performance of the thermometer segment is significantly better than that of the binary-weighted segment. In fact, each additional binary-weighted bit contributes a factor of two to the total harmonic distortion (THD) of the DAC, as shown in Figure 3.20 [32]. Therefore, the optimal segmentation occurs at the rightmost part of the horizontal line in the plot corresponding to the INL analog area constraint, which is still a minimum area design (constrained by INL requirements), meets the DNL specification, and exhibits the smallest THD.

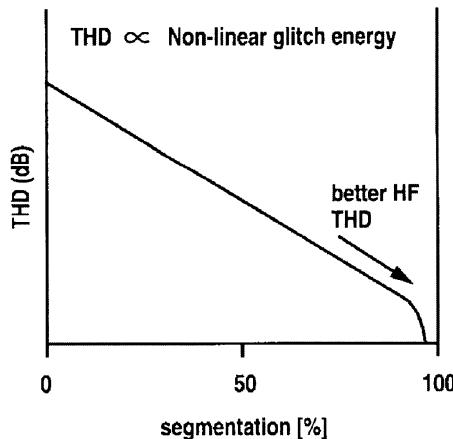


Figure 3.20. Total harmonic distortion (THD) versus segmentation for segmented current-steering DACs. Reprinted with permission [32].

3.5.2 Optimum segmentation based on spectral performance

For many communication applications it is more useful to use the Spurious Free Dynamic Range (SFDR) figure instead of the THD to characterize the spectral performance of the DAC. In the previous sub-section only the glitch energy contribution to the distortion has been considered. However, there are other dynamic effects of the converter that impact the spectral purity of the converted signals.

One important source of harmonic distortion, besides glitches, is input-dependent clock-jitter due to the simultaneous switching noise (SSN or dI/dt noise) generated in the binary to thermometer decoder and coupled to the clock-driver [37]. There is a second possible situation, where the clock-jitter originates from SSN coupled to the driver but generated in another digital circuitry integrated in the same chip than the DAC. In this later case, the clock-jitter is not correlated with the input signal and it produces an increase in the noise floor of the converter and hence a reduction of the SNR.

The SSN found in the internal digital power rails produces transient variations of the on-chip power supply voltage. If this noise is assumed to be coupled to the power supply nodes of the on-chip clock driver¹, its delay becomes a function of the SSN amplitude. When SSN is not correlated with the input signal, the relevant SSN amplitude occurs at the sampling times corresponding to the rising edge of the clock signal. The clock-driver power supply voltage during the clock rising edge transition can be considered a Gaussian random process with mean equal to the DC power supply voltage and standard deviation proportional to the SSN amplitude standard deviation. The resulting clock signal considered along a large number of output samples exhibits a random phase noise or clock-jitter with the same standard deviation (σ_{jitter}) of the SSN induced delay variation of the clock-driver. The SNR of the output analog signal is reduced proportionally to the clock-jitter standard deviation, because the Gaussian phase-noise in the output samples can be modeled to produce a amplitude error through the mean of the slew rate of the output signal [38,39]. This time to amplitude conversion process makes the SNR reduction due to random clock-jitter input-frequency dependent:

$$SNR = \frac{1}{2\pi^2 f_{in}^2 \sigma_{jitter}^2} . \quad (3.38)$$

¹ SSN is always coupled to the power supply nodes of the clock driver up to some extent, either because they are common with other digital circuitry, or, even when there are dedicated power supply nodes, because substrate noise is coupled through the biasing taps.

When the dominant SSN affecting the clock-driver is generated by the input decoder, there exists a correlation between the amplitude of the SSN and the sample being converted. There is also a synchronization between the instant at which the noise is generated (the change in the input data vector) and the arrival of the input clock rising edge. This implies that the input decoder SSN affects the delay of the clock-driver at the same time every input sample. The SSN amplitude generated when the input sample changes is determined mostly by the number of bits that are different between the current sample and the previous one in the binary to thermometer decoder. The binary segment bits are normally connected to a dummy decoder consisting in a chain of small size inverters used to emulate the delay of the thermometer decoder. Therefore the SSN generated by the dummy decoder compared with the thermometer decoder is negligible. The worst case SSN is generated when the entire thermometer segment bits switch simultaneously from one sample to the next. For a sinusoidal waveform with a frequency f_{in} much smaller than the sampling frequency f_s , the worst case SSN corresponds to the mid code transition. The number of bits switching, and hence the SSN amplitude, is proportional to the number of thermometer bits, M , and the worst case SSN occurs every $1/2f_{in}$ seconds. Consider for simplicity that all the SSN generated by the decoder can be represented by this worst case SSN. When this SSN is coupled to the clock-driver its delay suffers a deterministic change every $[f_s/2f_{in}]$ samples (where $[\cdot]$ denotes the closest integer). Moreover, this deterministic phase error or clock-jitter occurs when the slope of the sinusoidal is at its maximum. Therefore, the amplitude error due to this phase error is large. The time-to-amplitude error conversion is performed through a nonlinear function (the time derivative of the sinusoidal waveform). The error and the sinusoidal to which it is superimposed have a different frequency and the nonlinear operation produces intermodulation products that give rise to harmonic distortion at the output signal. The clock signal phase error (actually a delay error) due to SSN can be modeled by means of a deterministic impulse train signal. Using this model, the output signal including the amplitude error ($s_{dist}(t)$) produced by this SSN induced deterministic clock-jitter is expressed by:

$$s_{dist}(t) = A \sin(2\pi f_{in} t) + A 2\pi f_{in} \cos(2\pi f_{in} t) \cdot k \cdot M \cdot \sum_{i=-\infty}^{\infty} \delta\left(t - i \frac{1}{2f_{in}}\right), \quad (3.39)$$

where A is the signal amplitude and k is a parameter that accounts for the linear proportionality between the number of switching thermometer segment input bits and the delay change produced in the clock-driver (in seconds/bits units). The second term of eq. (3.39) is univocally described by a train of frequency

components, in which the effect of the third harmonic is dominant [40]. Considering only this third harmonic, the distorted output signal is expressed by:

$$s_{dist}(t) \triangleq A\sin(2\pi f_{in}t) + A \cdot k \cdot M \cdot 4\pi^2 f_{in}^2 \sin(6\pi f_{in}t), \quad (3.40)$$

From this expression, the SFDR is finally given by:

$$SFDR = \frac{1}{16\pi^4 k^2 M^2 f_{in}^2}, \quad (3.41)$$

The previous analysis is valid provided that the SSN is dominated by the mid code transition and that the input samples behave as an ordered sequence. This ordered sequence must include all the possible 2^N input codes inside each sinusoidal period (repeating some of them in the slow changing parts of the sinusoidal waveform around $\pi/2$ and $-\pi/2$). To assure that the sequence of all the 2^N codes appear at the input of the DAC it must be verified that in the worst case two samples correspond to two consecutive quantization levels. This occurs at the crossover point of the sinusoidal waveform where the slope is the highest possible, this is:

$$\frac{1 \text{ LSB}}{T_s} = \frac{2A/\cancel{2^N}}{\cancel{1/f_s}} \geq 2\pi f_{in} A \Leftrightarrow \frac{f_{in}}{f_s} \leq \frac{1}{2^N \pi}, \quad (3.42)$$

Actually only the bits connected to the thermometer decoder have to be taken into account. The sequence that needs to be considered is that containing the M thermometer bits, therefore the condition of eq. (3.42) is transformed into

$$\frac{f_{in}}{f_s} < \frac{1}{2^M \pi}, \quad (3.43)$$

that must be verified to apply the previous derivation of the SFDR.

When the condition of eq. (3.43) is far from being fulfilled, this is, when $f_{in}/f_s \gg 1/2^M \pi$, the sequence of samples appearing at the input of the thermometer decoder have a number of switching bits from sample to sample that can be considered a random process. In this case, the SSN amplitude generated in the decoder is not correlated with the input sequence, and the only appreciable effect

of this type of noise in the DAC spectral performance is a reduction of the SNR, as explained before.

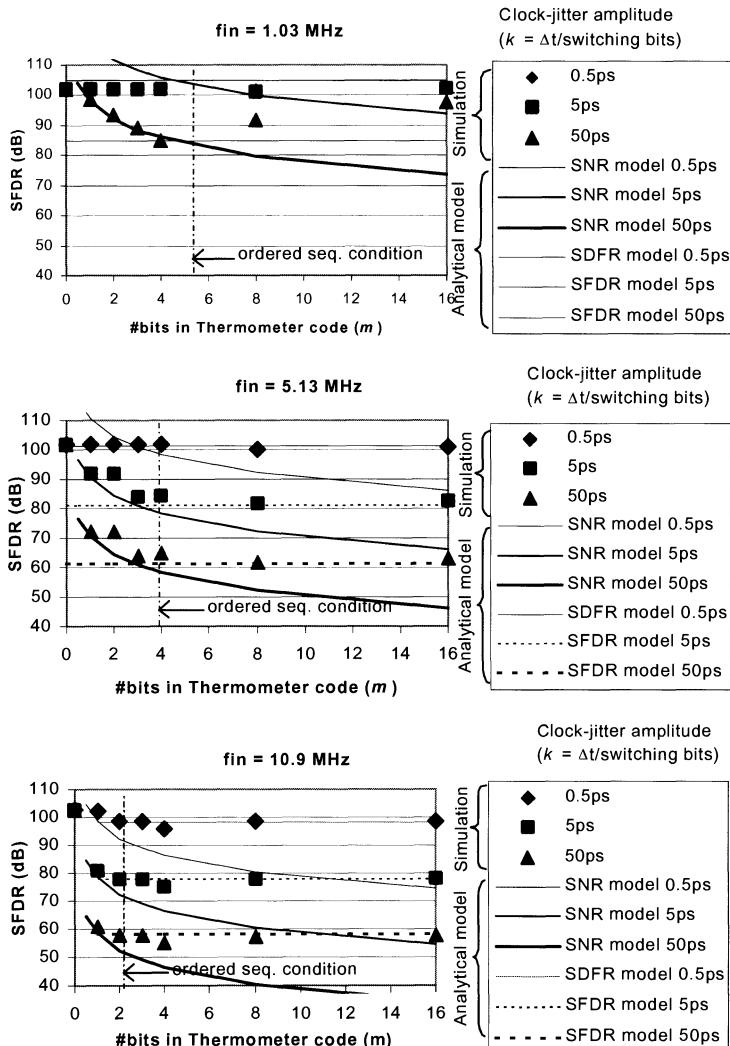


Figure 3.21. SFDR versus segmentation trade-off for a 16 bit current-steering DAC with a sampling frequency of 105 MHz, for different input frequencies and k values.

Now consider what occurs when increasing the number of thermometer bits M from 0 to N . First, for smaller M the condition of eq. (3.42) will be verified for a wide range of input frequencies. The SFDR due to clock-jitter in this region of small segmentation corresponds to eq. (3.41) and becomes smaller as M is increased. For a fixed frequency, when M is increased over the limit set by eq. (3.42), the clock-jitter can be considered as a random process and the SFDR measure collapses to the SNR expressed by eq. (3.38). The dependence of the SFDR on the segmentation ratio is illustrated in Figure 3.21 where the simulation results, obtained using a high level model of a 16-bit segmented current-steering DAC, are compared with the equations presented in this sub section for three different input signal frequencies.

To conclude this subsection note that, on one hand, it was shown in the previous sub section how the THD due to glitches is reduced when the segmentation of the DAC is increased (see Figure 3.20). On the other hand, it has been demonstrated in the previous paragraphs how the harmonic distortion increases when the segmentation is increased due to SSN and clock-jitter effects, as shown in Figure 3.21. Combining these two effects, an optimum segmentation can be found where the harmonic distortion due to both causes, glitches and clock-jitter, is minimized. This is illustrated in Figure 3.22.

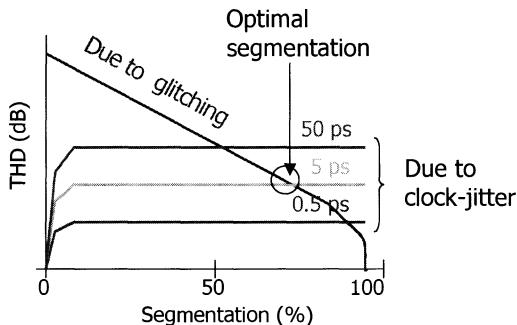


Figure 3.22. Optimum segmentation based on the spectral performance of a current-steering DAC.

3.6 PRACTICAL DESIGN OF CURRENT-STEERING D/A CONVERTERS

At architectural level, the decisions to be taken depend on the resolution and the static and dynamic specifications for the converter. The optimum segmentation was discussed in the previous section. Regarding the structure of

the thermometer segment of the DAC, for medium resolution converters (up to 10 bits) a row & column decoding structure and 1D switching sequences to compensate for systematic errors in the unary current cell array may be used. For higher resolutions, a single decoder and 2D hierarchical switching sequences should be used, as discussed in section 3.4. After the optimum architecture has been chosen, the design process continues by sizing the transistors of the basic building blocks of the LSB current cell. Then, each other current cell block is sized accordingly. The impact of the final layout phase is crucial. First, the arrays should be placed and routed according to the switching sequences, which for complex D/A structures is a cumbersome task. Next, special care has to be taken in the interconnection between all the large blocks of the converter and the isolation between noisy digital parts and sensitive analog signals. Other design issues to avoid unwanted couplings should be addressed at this stage as well, and will be addressed in the last part of this section.

3.6.1 Circuit sizing procedure

The goal of the circuit sizing procedure is to find the sizes and biasing voltages for the transistors of the LSB current cell circuit. That current cell consists, as presented in section 4.3, of a current source transistor (CS), two complementary switch transistors (SW) and, possibly, a cascode transistor (CAS) used to increase the output impedance of the current source (see Figures 4.9 and 4.10). During the sizing procedure, the following constraints are considered:

- Minimize area
- Minimize power consumption
- Maximize output impedance
- Minimize settling time
- Minimize glitches
- Guarantee a good matching and performance with a high parametric yield.

In most applications, it is important to maximize the output voltage swing, that for a CMOS current-steering converter is equal to the full scale output current multiplied by the output load. The full-scale output current on its own determines the analog section power consumption. The output voltage swing, the load value and the resolution are input specifications of the DAC. They determine the current of each current source.

3.6.1.1 Current source cell transistors

First, the case of a current cell consisting of a CS transistor and two complementary SW transistors will be addressed. The minimum area of the CS

transistor is determined by the mismatching parameters, the static performance requirements, the current value and the biasing voltage, as given by:

$$\begin{aligned} W_{CS}^2 &= \frac{I}{2K' \left(\frac{\sigma(I)}{I} \right)^2} \left[\frac{A_\beta^2}{(V_{gs} - V_T)_{CS}^2} + \frac{4A_{VT}^2}{(V_{gs} - V_T)_{CS}^4} \right] \\ L_{CS}^2 &= \frac{K'}{2I \left(\frac{\sigma(I)}{I} \right)^2} \left[A_\beta^2 (V_{gs} - V_T)_{CS}^2 + 4A_{VT}^2 \right] \end{aligned} \quad (3.44)$$

where the relative accuracy is related to the INL or DNL requirement (the most restrictive of them) by eq. (3.37). Therefore, for the CS transistor there is only one degree of freedom left to optimize the remaining constraints (output impedance, settling time, etc.).

The two complementary SW transistors are used to steer the current to one of the two complementary outputs of the current cell. In the ON state they are biased as a cascode transistor, in order to increase the output impedance. The ON gate voltage of the SW transistors determines the biasing of these transistors when acting as a cascode stage. Their aspect ratio and current will determine their transconductance, which is one of the parameters affecting the settling time. The SW transistors are usually minimum length transistors to improve the speed. Therefore, only their width and bias voltage have to be determined. However, both are related through the saturation current expression.

The two bias voltages of the CS and SW transistors (the SW ON voltage) are related, as both transistors must operate in saturation and the sum of their two drain to source voltages is limited by the maximum output swing. The condition for the bias (gate voltages) that guarantees that they operate in saturation is:

$$\underbrace{V_{OD}^{CS} + V_{OD}^{SW} + V_T^{SW}}_{V_{gSW}^{\min}} < V_{gSW} < \underbrace{V_{DD} - \Delta V_o^{\max} + V_T^{SW}}_{V_{gSW}^{\max}}, \quad (3.45)$$

where ΔV_o^{\max} is the maximum output voltage swing, V_{OD}^{CS} is the CS transistor overdrive voltage (gate-source voltage minus threshold voltage), and V_{OD}^{SW} is the SW transistor overdrive voltage. According to the minimum area requirement, the maximum possible overdrive voltage for the two transistors should be used since, for the same current, this gives the smaller aspect ratio. A solution exists for eq. (3.45) if and only if the left-hand part is smaller than the right-hand part. Thus the following condition is derived:

$$V_{OD}^{CS} + V_{OD}^{SW} \leq V_{DD} - \Delta V_o^{\max}, \quad (3.46)$$

from which the minimum area will be found when the left part is equal to the right part. These two last expressions relate the SW and the CS transistors bias voltages in such a way that if one of them is fixed, the other one is derived using eq. (3.46). The CS transistor is the larger of the two, so its overdrive voltage is always fixed to the highest possible value fulfilling eq. (3.46). By doing this, the overdrive voltage of the SW transistors is found just at the limit between the triode and the saturation regions. In order to have some margin to allow for process fluctuations and still have the two transistors in saturation, an arbitrary safety margin of some hundreds of milivolts may be introduced in eq. (3.46) [22,41].

Now, the previous expression will be used to find the optimum sizes and bias voltages. The optimization goal is the minimization of the settling time. In the circuit of Figure 4.9 there are two poles that determine the switching speed of the current cell. The first pole is due to the output load and the parasitic capacitance at the drain of the switch transistor (that will increase with its width). The second pole, due to the internal node, has contributions of both the parasitic capacitance of the CS drain and the SW source, and depends on the SW transistor small signal trasconductance and body effect parameters, as presented in [22]:

$$\begin{aligned} p_1 &\approx \frac{1}{2\pi R_L (C_L + C_{drainot}^{SW})} \\ p_2 &\approx \frac{g_m^{SW} + g_{mb}^{SW}}{2\pi C_{drainot}^{CS}}, \end{aligned} \quad (3.47)$$

All the small signal parameters and parasitic capacitances determining the poles are related between themselves through eqs. (3.44)-(3.46). The only degree of freedom is one of the two overdrive voltages (either that of the CS or the SW). Therefore, it is possible to plot the frequency of the two poles against this degree of freedom, as shown in Figure 3.23 for the particular converter presented in [22].

The internal pole p_2 can be moved to higher frequencies by increasing g_m^{SW} or by decreasing $C_{drainot}^{CS}$ restricted to the constraint set by the required accuracy specification given by eq. (3.44). To increase g_m^{SW} , for a fixed current, it is necessary to increase the SW transistor aspect ratio $(W/L)^{SW}$. This increases the parasitic capacitance $C_{drainot}^{SW}$ (for a minimum gate length), reducing the output pole p_1 frequency. The optimum sizes and bias voltages are found when the two pole frequencies are coincident, and this, in the figure, determines an optimum V_{OD}^{CS} , which then determines the remaining CS and SW transistors sizes and voltages.

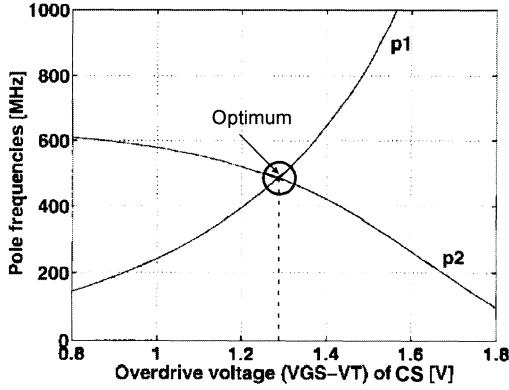


Figure 3.23. Pole locations as function of the overdrive voltage of the current source transistor [22].

There is a complementary approach to optimally size the current cell circuit that eliminates the need to introduce an arbitrary safety margin in the saturation condition of eq. (3.46) and that takes into account the maximization of the output impedance. First, it is easy to find that, as the CS gate voltage is intrinsically determined by its overdrive voltage, the maximum of the DC output impedance occurs when the SW gate voltage is:

$$V_{gSW} = V_T^{SW} + \frac{1}{2} \left(V_{DD} - \Delta V_o^{\max} + V_{OD}^{CS} + V_{OD}^{SW} \right) = \frac{V_{gSW}^{\max} + V_{gSW}^{\min}}{2}. \quad (3.48)$$

In second place, process variation errors affecting the SW transistors can be included in the analytical formulation of the saturation condition in a similar way as with the mismatching errors for the CS transistor. The two bounds of eq. (3.45) are then modeled by means of a Gaussian distribution. The variance of the upper bound is found by, first, expressing ΔV_o^{\max} as a function of the LSB current i_0 and the load resistance R_L , and, second, taking partial derivatives of all the terms:

$$\begin{aligned} \sigma_{V_{gSW}^{\max}}^2 &\approx \left(\frac{\partial V_{gSW}^{\max}}{\partial i_0} \right)^2 \sigma_{i_0}^2 + \left(\frac{\partial V_{gSW}^{\max}}{\partial V_T^{SW}} \right)^2 \sigma_{V_T^{SW}}^2 + \left(\frac{\partial V_{gSW}^{\max}}{\partial R_L} \right)^2 \sigma_{R_L}^2 = \\ &= \frac{A_{VT}^2}{W_{SW} L_{SW}} + \Delta V_o^{\max 2} \left(\frac{\sigma_{i_0}^2}{i_0^2} + \frac{\sigma_{R_L}^2}{R_L^2} \right) \end{aligned} \quad (3.49)$$

Similarly, the variance of the lower bound yields:

$$\sigma_{V_{gSW}^{\min}}^2 \approx \frac{A_{VT}^2}{W_{CS}L_{CS}} + \frac{A_{VT}^2}{W_{SW}L_{SW}} + \frac{V_{OD}^{SW^2}}{4} \left(\frac{\sigma_{i_0}^2}{i_0^2} + \frac{A_\beta^2}{W_{SW}L_{SW}K'^2} \right). \quad (3.50)$$

To find an appropriate value for the SW gate voltage, the upper bound must be larger than the lower bound in a given percentage of the cases expressed by the *yield_SW*. If this is accomplished, the optimum of the SW gate voltage found in eq. (3.48) has to verify that:

$$\begin{aligned} p\left(\left[V_{gSW}^{\max} - \frac{V_{gSW}^{\max} + V_{gSW}^{\min}}{2}\right] > 0\right) &\geq \text{yield_SW} \text{ and} \\ p\left(\left[\frac{V_{gSW}^{\max} + V_{gSW}^{\min}}{2} - V_{gSW}^{\min}\right] > 0\right) &\geq \text{yield_SW} \end{aligned}, \quad (3.51)$$

which can be expressed also as:

$$\frac{V_{gSW}^{\max} - V_{gSW}^{\min}}{2} \geq S \max\left[\sigma_{V_{gSW}^{\max}}, \sigma_{V_{gSW}^{\min}}\right], \quad (3.52)$$

where S is the number of standard deviations of the Gaussian distribution that are needed to fulfill eq. (3.51): $S = \text{inv_norm}(\text{yield_SW})$. The *yield_SW* is related to the INL yield by the following formula:

$$\text{yield} = \text{yield_SW}^2, \quad (3.53)$$

that expresses the fact that for every sample considered in the INL statistics all the two switches of the DAC LSB current cell (the worst case because it has the smallest transistors) must individually verify the saturation condition with a probability expressed by *yield_SW* (all the SW transistors of different current cells are considered to have independent errors with the same probability distribution).

The condition of eq. (3.52) is interpreted as a safety region on the plane formed by all the possible combinations of CS and SW overdrive voltages. This condition replaces eq. (3.46) in the sizing procedure and avoids the introduction of an arbitrary safety margin because it implicitly takes into account process variations. In order to illustrate this complementary sizing approach, the expressions presented above have been programmed into a MATLAB script. By providing some inputs to the script, such as the number of bits of the converter, the segmentation, power supply voltage, maximum output swing, load resistance,

and the technology parameters, the script can be used to explore the design space and find an optimum for different particular goals.

Figure 3.24 shows two plots for the same design example consisting of a 12-bits DAC with $B = 4$, $M = 8$, $V_{DD} = 3.3$ V, $\Delta V_o^{\max} = 1$ V, and $R_L = 50 \Omega$. In the top plot the x - y plane is the available design space corresponding to the CS and SW transistor areas (that are determined by their overdrive voltages). The z -axis

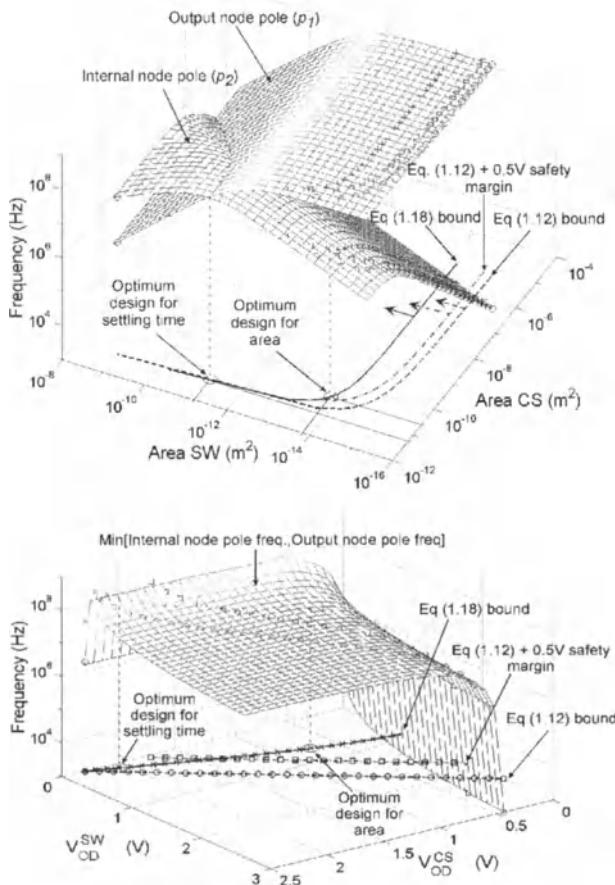


Figure 3.24. Optimization graphs for a current source and switch transistors current cell sizing in a 12-bit current-steering DAC.

shows the frequency position of the two poles. The curves within the x - y plane correspond to three different saturation conditions: eq. (3.46), eq. (3.46) plus a 0.5 V safety margin, and eq. (3.52). The graph on the bottom shows the minimum of the two poles frequency in the z -axis and the CS and SW overdrive voltages in the x - y plane. In that graph it is more evident that using eq. (3.46) with an arbitrary safety margin can lead either to too pessimistic or too optimistic designs. For example, the optimum design point for settling time would produce larger CS transistors using eq. (3.46) than using eq. (3.52). In the other hand, the optimum design point for minimum area found using eq. (3.46) would be out of the realistic safety design region found by eq. (3.52), indicating that for that case, the 0.5 V safety margin is too optimistic.

The other usual current cell circuit topology includes an additional cascode transistor (CAS) to increase the output impedance. In this case, the same procedure presented above can be used to simultaneously size all the transistors of the circuit. This new transistor in series contributes with an additional saturation condition to be fulfilled. The optimum of the DC output impedance is found when the SW and the CAS gates voltages verify the following expressions:

$$\begin{aligned} V_{g_{CAS}} &= V_T^{CAS} + \frac{1}{3} \left[V_{DD} - \Delta V_o^{\max} + 2V_{OD}^{CS} + 2V_{OD}^{CAS} - V_{OD}^{SW} \right] \\ V_{g_{SW}} &= V_T^{SW} + \frac{1}{3} \left[2(V_{DD} - \Delta V_o^{\max}) + V_{OD}^{CS} + V_{OD}^{CAS} + V_{OD}^{SW} \right], \end{aligned} \quad (3.54)$$

The CAS transistor introduces two new degrees of freedom in the design procedure: its overdrive voltage and its area (or channel length, as the aspect ratio is fixed by the overdrive voltage and the current). One of these degrees of freedom can be eliminated as discussed in [21]. In that work the output impedance at the Nyquist bandwidth is analyzed and it is obtained a relation between the CAS and SW transistor that maximizes this AC output impedance. Another possible criterion is to use a minimum area CAS transistor that also minimizes the parasitic capacitance at the source of the SW transistor, and therefore if the pole due to that capacitance is the dominating pole, the settling time is reduced. Once this degree of freedom is eliminated, there are three degrees of freedom left: the three overdrive voltages of the transistors. However, not all the possible combinations will guarantee that the three transistors operate in saturation including process variations in all the transistors. The safety region (now a 3D volume) that constraints the design space is found by applying the same Gaussian distribution model to the bounds of the SW and CAS transistors overdrive voltages, in a similar way to what was made previously only for the SW transistor:

$$\begin{aligned}
\sigma_{V_{gSW}^{\max}}^2 &\approx \frac{A_{VT}^2}{W_{SW}L_{SW}} + \Delta V_o^{\max 2} \left(\frac{\sigma_{i_0}^2}{i_0^2} + \frac{\sigma_{R_L}^2}{R_L} \right) \\
\sigma_{V_{gSW}^{\min}}^2 &\approx \frac{A_{VT}^2}{W_{SW}L_{SW}} + \frac{A_{VT}^2}{W_{CAS}L_{CAS}} + \frac{V_{OD}^{SW2}}{4} \left(\frac{\sigma_{i_0}^2}{i_0^2} + \frac{A_\beta^2}{W_{SW}L_{SW}K^{12}} \right) \\
\sigma_{V_{gCAS}^{\max}}^2 &\approx \frac{A_{VT}^2}{W_{SW}L_{SW}} + \frac{A_{VT}^2}{W_{CAS}L_{CAS}} + \frac{V_{OD}^{SW2}}{4} \left(\frac{\sigma_{i_0}^2}{i_0^2} + \frac{A_\beta^2}{W_{SW}L_{SW}K^{12}} \right). \\
\sigma_{V_{gCAS}^{\min}}^2 &\approx \frac{A_{VT}^2}{W_{CS}L_{CS}} + \frac{A_{VT}^2}{W_{CAS}L_{CAS}} + \frac{V_{OD}^{CAS2}}{4} \left(\frac{\sigma_{i_0}^2}{i_0^2} + \frac{A_\beta^2}{W_{CAS}L_{CAS}K^{12}} \right)
\end{aligned} \tag{3.55}$$

Now the saturation condition is double:

$$\begin{aligned}
\frac{V_{gSW}^{\max} - V_{gSW}^{\min}}{2} &\geq S \max \left[\sigma_{V_{gSW}^{\max}}, \sigma_{V_{gSW}^{\min}} \right] \\
\frac{V_{gCAS}^{\max} - V_{gCAS}^{\min}}{2} &\geq S \max \left[\sigma_{V_{gCAS}^{\max}}, \sigma_{V_{gCAS}^{\min}} \right].
\end{aligned} \tag{3.56}$$

These two expressions correspond to a 3D surface that limits the design space volume. Figure 3.25 shows a graph that plots this surface in the 3D design space, where each axis corresponds to one of the CS, SE or CAS overdrive voltages for the same 12-bit DAC design example previously presented. The bound expressed

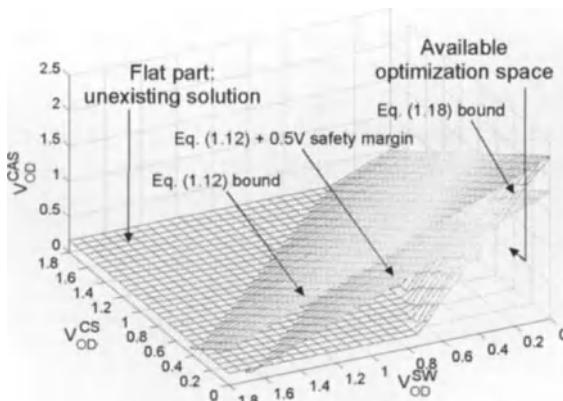


Figure 3.25. Design space for the switch + current source + cascode current cell topology 12-bit current-steering DAC.

by eq. (3.56) is plotted along with the bounds obtained using eq. (3.46) with and without the additional safety margin of 0.5 V for comparison purposes. In this graph, a 4th dimension would be required to represent the optimization parameter (for example, the pole frequencies) against the design space. However, the MATLAB script can be used to find an optimum point inside the constrained design space for any optimization criteria.

3.6.1.2 Latch and switch driver block

Three important issues that have been identified to cause dynamic limitations in current-steering D/A converters are [33]:

- Imperfect synchronization of the control signal at the switches.
- Drain voltage variations of the current-source transistors.
- Coupling of the control signal through the gate to drain capacitance of the switches to the output

These problems can be solved with a properly designed latch placed just before the switch transistors. It is very important that during the switching of the two complementary switch transistors the current flowing through the two branches always sum the constant current of the CS transistor. If during some period of time both SW transistors are OFF the CS current will make its drain node voltage to vary, producing a glitch at the output. To reduce these glitches the two complementary SW gate voltage waveforms must be generated in such a way that when one of the two SW transistors is turned OFF the other is completely turned ON [42]. The two complementary gate voltage waveforms that control the switch operation must have a non-symmetrical crossover. The switch gate ON voltage is determined during the sizing procedure explained in the previous sub section and is smaller than V_{DD} . The gate OFF voltage can be also larger than Gnd. The reduction of the switch gate voltage swing helps in reducing the glitches generated at the output node due to clock feedthrough. A driver

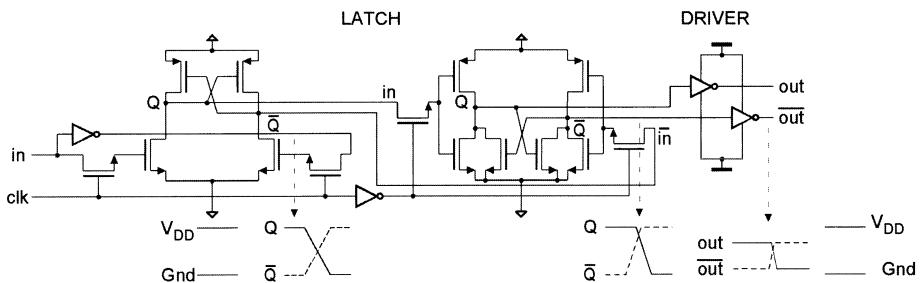


Figure 3.26. Latch and driver circuit.

connected to the outputs of the latch provides these voltage levels. One example of latch + driver block and its corresponding waveforms is shown in Figure 3.26, which is a modification of the circuit proposed in [33].

3.6.1.3 Biasing circuits

There are several bias voltages that must be generated and distributed to all the current source cells of the DAC. The most important one is the bias voltage of the current source transistor. This voltage is usually generated on-chip. For a current-steering segmented DAC with M thermometer bits, $2^M - 1$ unary current sources are required. The unary current sources are organized in an array, as explained in the previous sections and there is always an empty space because the number of unary current sources is odd. The empty space is used to place a unary current source transistor in diode connection that generates the current source gate bias voltage and mirrors it to all the other current source transistors of the DAC. The other bias voltages are the gate voltage of the cascode transistors and the ON and OFF voltages for the switches that usually are generated off-chip for test chips. All the biasing voltages in commercial applications are generated with on-chip biasing circuits designed following standard techniques [43].

3.6.2 Layout issues

In this subsection, it is included a list of important considerations to be taken into account when designing the layout of the different blocks and the top level floorplan of the converter.

3.6.2.1 Current cell array

It is recommended to separate the current source transistors from the switches and other inherently noisy circuitry and place them in separate locations [44]. The current source array is composed only of current source transistors and, possibly, cascode transistors. The structure should be as regular as possible, especially if a single decoder and a 2D switching sequence are used, to facilitate routing. Each current cell should have the same metal layers and structures on top of it. Assuming a 3-metal technology, a proper metal layer to signal assignment follows:

- Metal 1 for ground + biasing, in one direction.
- Metal 2 in the perpendicular direction to contact the drains of the current source transistors.

- Metal 3 to connect together the drains of the transistors that belong to the same current cell and to connect this node to the corresponding switches that are placed in a separate array.

This interconnection structure is routed on top of the current cell array. The ground lines should be wide enough to carry the full-scale output current. In order to minimize edge effects two or three rows of dummy cells should be placed in the edges of the array.

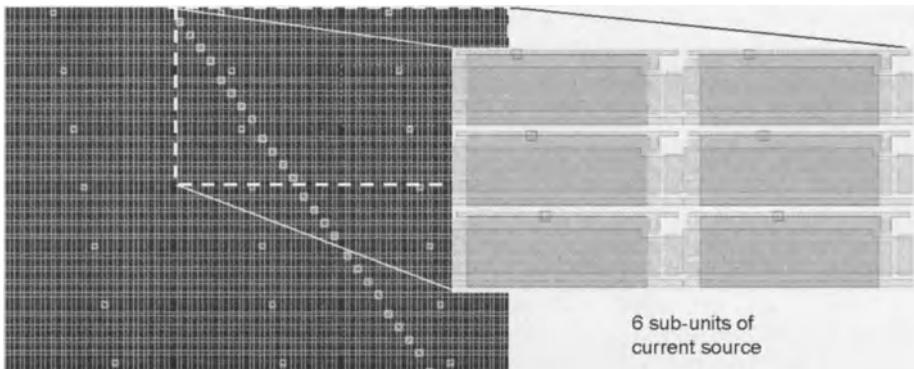


Figure 3.27. Layout of a portion of the current source transistors array and a detail showing the transistors and the metal 1 interconnections that distribute the ground and the bias throughout the array.

3.6.2.2 *Latch, driver and switch arrays*

For row-column decoder architectures, the latches can be placed in the edges of the current source array just after the decoders. The switch transistors are included inside each cell of the array because a local decoding logic is necessary for each current source. This requirement increases the total current source array area, aggravating the systematic error effects, and places digital circuitry close to the sensitive analog transistors. The main benefit of these architectures is that the routing is very straightforward and 1D switching sequences to compensate for systematic errors are simple to implement.

Other switching sequence techniques require the use of a single thermometer decoder placed on top of the array. In this case, the latches, drivers and switch transistors, as well as the current source transistors are placed in two separate arrays. This allows to separate digital and analog circuits. The switches for the binary bits should be proportionally scaled to match the weight of their current source values. The latch+driver circuit is the same for the binary and for the

unary sources. Therefore, as the unary current cell switches are wider than the binary current cell switches, dummy switches are placed at the output of the binary latch+drivers to match load and delay in all the latch+driver blocks. Usually the binary and unary latch+driver and switches are placed together in the same array. It is useful to locally use there a centroid distribution as in the current source transistor array (see Figure 3.28).

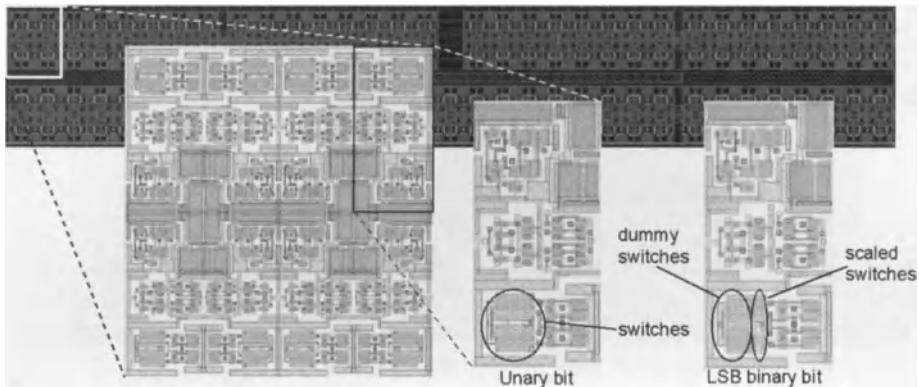


Figure 3.28. Layout of the latch+driver+switches array with details of the centroid distribution and unary and binary block examples.

The clock signal has to be distributed to all the latch clock input ports reducing any skew. A clock tree is a good approach with a well sized on-chip clock driver in its root (see top layout of Figure 3.29). The clock signal should be routed in such a way that the crossing with other sensitive analog lines in the latch+driver+switches array is avoided. One of the most critical analog interconnections are the two complementary output lines that collect the output current of the current sources and connect them to the output of the DAC. These lines should be wide enough, being a good recommendation to route shielding lines surrounding them, as shown in the bottom layout of Figure 3.29.

3.6.2.3 Thermometer decoder

This combinational circuit can be automatically synthesized using a digital standard-cell library and then automatically placed and routed. Timing driven place and route is recommended for high-speed DAC. However, if the segmentation ratio is large it is possible that this block becomes the timing bottleneck of the converter. In this case it can be either pipelined [30] or implemented using custom logic cells and custom placed and routed to optimize the delay [33]. If a pipeline structure is used it is recommended to use a separated clock driver for the latch+driver+switches block than for the decoder, to

minimize noise coupling between these two blocks that may induce clock-jitter and distortion, as explained in section 3.5.

A dummy decoder is used for the binary bits to equalize the latency with the thermometer decoder.

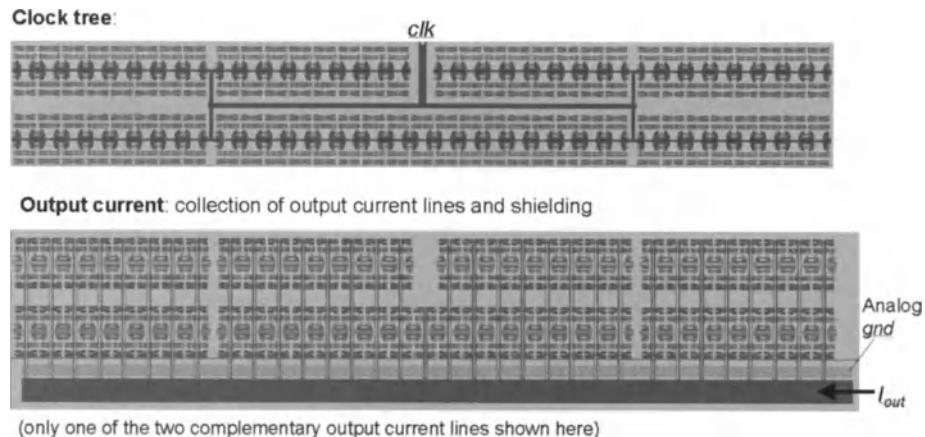


Figure 3.29. Layout of the latch+driver+switches array with details of the routing of the clock signal and one of the two complementary output current lines.

3.6.2.4 Top level floorplan

Being the DAC a mixed signal circuit, it is therefore crucial to isolate the analog parts from the digital parts and to avoid the coupling of digital noise to the analog circuits [45]. Moreover, there are two different types of digital circuits: the input decoders and the clock driver and latches. The first type of circuit can tolerate larger levels of switching noise (dI/dt noise) but the second type should have a power supply as stable as possible. As explained in 3.5 any noise at the latch+driver+switches clock driver produces phase errors in the clock signal and this has an important impact over the DAC spectral performance (reducing the SNR and even causing distortion). The best solution is to use separated power supply on-chip distributions for the input decoders and for the latch+driver+switches array and clock driver. Using separate power supply package pins, if possible, helps in avoiding coupling as well. The analog section should have a dedicated power supply distribution and package power supply pins, also.

Even though separated on-chip power supply distributions are used for the different sections of the decoder, the common substrate couples the noise from one section to the others. There are two ways of addressing this issue:

- Reduce the dI/dt noise generation in the digital sections by using multiple package pins (reduce inductance) and on-chip decoupling capacitors.
- Reduce substrate coupling by placing analog blocks far from the noisiest blocks (i.e., input decoder and clock-driver) and use guard rings surrounding the sensitive analog sections. The rings must be properly biased with dedicated substrate taps and package pins.

All these techniques are specially important for high-speed high-accuracy D/A converters where the fast switching times and the complexity of the input decoders will respectively produce large individual dI/dt noise for each digital cell and a large amount of simultaneous switching noise (SSN).

3.6.3 A high-speed high-accuracy current-steering D/A converter design example

A high-performance 12-bit D/A converter is presented in this subsection to illustrate the design issues previously presented. The DAC was originally designed to study the effects of dI/dt noise on the converter performance. The converter has a segmented architecture with 4 binary bits and 8 thermometer bits. The top level structure of the DAC is shown in Figure 3.30. It has a single decoder because a 2D hierarchical switching sequence is used in the current source array to compensate for systematic errors.

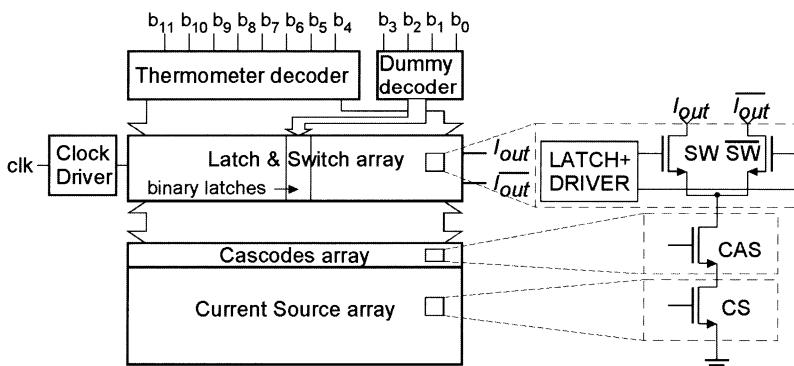


Figure 3.30. Floorplan of the 12-bit current-steering D/A converter design example.

The thermometer decoder has been synthesized using commercial synthesis tools from a VHDL functional description based in the decoder presented in [30]. A timing driven synthesis and place & route tool has been used, including estimation of the layout parasitics. The worst case estimated delay is 2.20 ns. Dummy cells for the binary bits are included in the decoder block having approximately the same delay that the average thermometer decoder path delay.

The basic current cell is composed of a NMOS current source transistor, a NMOS cascode transistor and two complementary NMOS switch transistors, driven with reduced swing control signals generated by the latch+driver block shown in Figure 3.26 and Figure 3.28. The unary current source transistors are placed in a 256×256 array and their latch+driver+switches are placed in a separate array on top of the current source array. The binary current source transistors are placed in four columns symmetrically inserted around the center of the unary current source array, while their latch+driver+switches are placed in the middle of the unary latch+driver+switches array. All the cascode transistors are placed in a small array on top of the current source array, as shown in Figure 3.31.

The unary current source transistors are divided in 16 parallel connected NMOS transistors that are placed across the array following a double centroid distribution [30]. Therefore, the 256×256 array is actually divided in 16 sections. Each section has one of the 16 components of a unary current cell and these units are distributed using the optimal 2D hierarchical switching sequence presented in [31]. A detail of the current source array including these basic transistors is shown in Figure 3.27. The binary current source transistors are also implemented by connecting in parallel or in series the required number of these basic transistors. The interconnection between the individual transistors and the routing with the latch+driver+switches array is done following the guidelines given in [30]. The place and route of the two arrays is done by the CAD tool from a physical description file that is automatically generated using a C routine and using full custom layouts of the basic building blocks.

The complete layout of the converter is shown in Figure 3.31 for illustration purposes. It has been designed using a digital $0.35 \mu\text{m}$ CMOS process with 1 poly and 3 metal layers. The power supply is 3.3 V volts and has a full scale current of 20 mA over 50 ohms loads. It occupies an active area (excluding I/O pads) of $1.65 \times 1.08 \text{ mm}^2$ and a simulated maximum speed of 400 Msamples/s. Multiple package pins have been used for each power supply (decoder, clock driver & latch array and analog section) to reduce dI/dt noise. Two package pins are used for each of the complementary current outputs to reduce the series parasitic inductance. On-chip decoupling capacitors are also included in any empty space of the layout and also inside the empty spaces of the latch+driver+switches array. A stabilizing on-chip capacitor is also included

between the on-chip reference voltage of the current-source array bias generator and the analog ground.

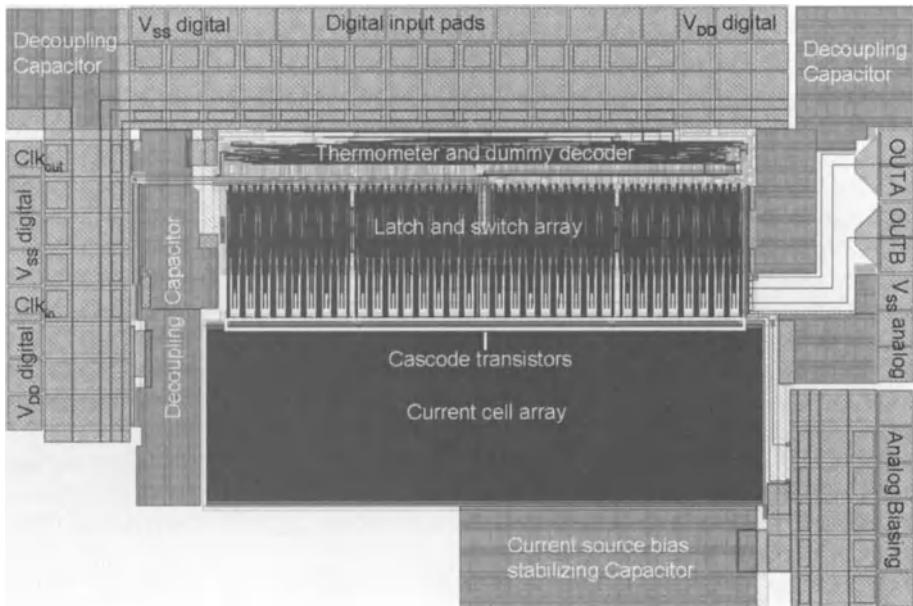


Figure 3.31. Top level layout of the 12-bit current-steering D/A converter design example.

3.7 References

- [1] H. Samueli, "Broadband Communications ICs: Enabling high-bandwidth connectivity in the home and office", in *Proc. IEEE 1999 ISSCC*, Feb. 1999, pp. 26-30.
- [2] A.R. Bugeja, B.S. Song, P.L. Rakers and S.F. Gillig , "A 14-b, 100-MS/s CMOS DAC designed for spectral performance", *IEEE Journal of Solid-State Circuits*, Vol. 34, No. 12, December 1999, pp. 1719-1731.
- [3] Randall L. Geiger, Philip E. Allen, Noel R. Strader, *VLSI Design techniques for analog and digital circuits*, McGraw-Hill International Editions, 1990.
- [4] Mikael Gustavsson, J. Jacob Wikner, and Nianxiong Nick Tan, *CMOS Data Converters for Communications*, Kluwer Academic Publishers, 2000.
- [5] K. Bult and G. J. M. Geelen, "An inherently linear and compact MOST-only current division technique," *IEEE J. Solid-State Circuits*, vol. 27, Dec. 1992, pp. 1730–1735.
- [6] R.S. Soin, F. Maloberti, and J. Franca, *Analogue-Digital ASICs*, Peter Peregrinus Ltd., 1991.

- [7] Paul Hendriks, "Specifying Communication DACs", *IEEE Spectrum*, July 1997, pp. 58-69.
- [8] Naoki Kumazawa, Noriyuki Fukushima, Naomi Ono, Hobuhiro Sakamoto, "An 8 bit 150 MHz converter with 2 Vp-p wide range output", *IEEE Symposium on VLSI Circuits*, 1990, pp. 55-56.
- [9] Tsutomu Kamoto, Yukiro Akazawa, Mitsuru Shinagawa, "An 8-bit 2-ns Monolithic DAC", *IEEE Journal of Solid-State Circuits*, Vol. 23, No. 1, February 1998, pp. 142-146.
- [10] Lauri Sumanen, Mikko Waltari, Kari Halonen, "A 10-bit High-Speed Low-Power CMOS D/A Converter in 0.2mm²", *IEEE Intl. Conference on Circuits and Systems*, 1998, pp 15-18 vol. 1.
- [11] Nicholas van Bavel, "A 325 MHz 3.3V 10-Bit CMOS D/A Converter Core With Novel Latching Driver Circuit", *IEEE Custom Integrated Circuits Conference*, 1998, pp. 245-247.
- [12] Jin-Park, Seung-Chul Lee, and Seung-Hoon Lee, "3V 10b 70MHz CMOS D/A converter for video applications", *Electronic Letters*, Vol. 35, No. 24, November 1999, pp. 2071-2073.
- [13] Seung-Chul Lee, Jin Park, Jin-Sik Yoon, Jung-Hee Song, and Seung-Hoon Lee, "A 3 V 100 MS/s Digital-to-Analog Converter for Cable Modem Applications", *IEEE Transactions on Consumer Electronics*, Vol. 46, No: 4 , November 2000, pp. 1043 -1047
- [14] C-H. Lin, K. Bult, "A 10b 250Msample/s CMOS DAC in 1mm²", *IEEE Intl. Solid-State Integrated Circuits Conference*, 1998, pp. 214-215.
- [15] KI-Hong Ryu, Kwang Sub Yoon, Hing Ki Min, "Design of a 3.3V 12bit CMOS D/A Converter with a high linearity", *Proceedings. 1998 Midwest Symposium on Circuits and Systems*, 1998, pp 538-541.
- [16] A. Marques, J. Bastos, A. Van den Bosch, J. Vandebussche, M. Steyaert, W. Sansen, "A 12b Accuracy 300Msample/s Update Rate CMOS DAC", *IEEE Intl. Solid-State Integrated Circuits Conference*, 1998, pp. 216-217.
- [17] A. Van den Bosch, M. Borremans, J. Vandebussche, G. Van der Plas, A. Marques, J. Bastos, M. Steyaert, G. Gielen, W. Sansen, "A 12 bit 200 MHz Low Glitch CMOS D/A Converter", *IEEE Custom Integrated Circuits Conference*, 1998, pp 249-252.
- [18] Bruce J. Tesch, and Juan C. Garcia, "A Low Glitch 14-b 100-MHz D/A Converter", *IEEE J. of Solid-State Circuits*, Vol. 32, No. 9, September 1997, pp. 1465-1469.
- [19] Alex R. Bugeja, Bang-Sup Song, "A Self-Trimming 14b 100Msamples/s CMOS DAC", *IEEE Intl. Solid-State Integrated Circuits Conference*, 2000, pp. 44-45.
- [20] Marcel J.M Pelgrom, Aad C. J. Duinmaijer, and Anton P.G. Welbers, "Matching Properties of MOS Transistors", *IEEE J. of Solid-State Circuits*, vol. 24, no. 5, 1989, pp. 1433-1440.
- [21] A. Van den Bosch, M. Steyaert, and W. Sansen, "SFDR-Bandwidth Limitations for High Speed High Resolution Current Steering CMOS D/A Converters",

- Proceedings of the IEEE International Conference on Electronics, Circuits and Systems (ICECS), Sept. 1999, pp 1193-1196.*
- [22] J. Bastos, A.M. Marques, M.S.J. Steyaert, W. Sansen, "A 12-Bit Intrinsic Accuracy High-Speed CMOS DAC", *IEEE J. of Solid-State Circuits*, Vol. 33, No. 12, 1998, pp. 1959-1969.
 - [23] Xavier Aragonès, José Luis González, and Antonio Rubio, *Analysis and Solutions for Switching Noise Coupling in Mixed-Signal ICs*. Kluwer Academic Publishers. Boston, 1999.
 - [24] Frank Herzl, and Behzad Razavi, "A Study of Oscillator Jitter Due to Supply and Substrate Noise", *IEEE Tr. on Circuits and Systems-II*, Vol. 46, No. 1, Jan. 1999, pp. 56-62.
 - [25] Patrik Larsson, "Resonance and Damping in CMOS Circuits with On-Chip Decoupling Capacitance", *IEEE Tr. on Circuits and Systems-I*, Vol. 45, No. 8, Aug. 1998, pp. 849-858.
 - [26] M. Orshansky, L. Milor, Ly Nguyen, G. Hill, Yeng Peng, and Chenming Hu, "Intra-gate CD variability and its impact on circuit performance", *Intl. Electron Devices Meeting 1999 Technical Digest*, 1999, pp. 479-482.
 - [27] B.E. Stine, V. Mehrotra, D.S. Boning, J.E. Cung, and D.J. Ciplickas, "A simulation methodology for assessing the impact of spatial/pattern dependent interconnect parameter variation on circuit performance", *Intl. Electron Devices Meeting 1997 Technical Digest*, 1997, pp. 132-136.
 - [28] J. M Fournier, P. Senn, "A 130-MHz 8-b CMOS Video DAC for HDTV Applications", *IEEE J. of Solid-State Circuits*, vol. 26, no. 7, 1991, pp. 1073-1076.
 - [29] Y. Nakamura, T. Miki, A. Maeda, H. Kondoh, N. Yazawa, "A 10-b 70-MS/s CMOS D/A converter", *IEEE J. of Solid-State Circuits*, vol. 26, no. 4, 1991, pp. 637-642.
 - [30] G. Van der Plas, J. Vandebussche, W. Sansen, M. Steyaert, G. Gielen, "A 14-b Intrinsic Accuracy Q²-Random Walk CMOS DAC", *IEEE J. of Solid-State Circuits*, vol. 34, no. 12, 1999, pp. 1708-1718.
 - [31] Yonghua Cong, and Randall L. Geiger, "Switching Sequence Optimization for Gradient Error Compensation in Thermometer-Decoded DAC Arrays", *IEEE Trans. on Circuits and Systems-II*, vol. 47, no. 7, 2000, pp. 585-595.
 - [32] Chi-Hung Lin, and Klaas Bult, "A 10-b, 500-Msample/s CMOS DAC in 0.6 mm²", *IEEE J. of Solid-State Circuits*, vol. 33, no. 12, 2000, pp. 1948-1958.
 - [33] A. Van den Bosch, M. A. F. Borremans, M. S. J. Steyaert, and W. Sansen, "A 10-bit 1-Gsample/s Nyquist Current-Steering CMOS D/A Converter", *IEEE J. of Solid-State Circuits*, vol. 36, no. 3, 2001, pp. 315-324.
 - [34] A. Van den Bosch, M. A. F. Borremans, M. S. J. Steyaert, and W. Sansen, "A 12b 500Msample/s Current-Steering CMOS D/A Converter", *Proc. of 2001 IEEE Intl. Solid-State Circuits Conference*, 2001, pp. 366-367.
 - [35] Alex R. Bugeja, and Bang-Sup Song, "A Self-Trimming 14.b 100-MS/s CMOS DAC", *IEEE J. of Solid-State Circuits*, vol. 35, no. 12, 2000, pp. 1841-1852.

- [36] Mika Tiilikainen, "A 1.8V 20mW 1mm^2 100MSample/s CMOS DAC", *Proc. of 26th European Solid-State Circuits Conference (ESSCIRC)*, 2000, pp.
- [37] J.L González, Eduard Alarcón, "Clock-Jitter Induced Distortion in High-Speed CMOS Switched-Current Segmented Digital-to-Analog Converters", *Proceedings of the IEEE Intl. Symposium on Circuits and Systems (ISCAS'01)*, Sydney, Australia, May 2001, pp. I-512-I-515.
- [38] Selim Saad Award, "Analysis of Accumulated Timing-Jitter in the Time Domain", *IEEE Tr. on Instrumentation and Measurement*, Vol. 47, No. 1, Feb. 1988, pp. 69-73.
- [39] Harau Kobayashi, Masanao Morimura, Kensuke Keboyashi, and Yoshitaka Onaya, "Aperture Jitter Effects in Wideband Sampling Systems", *Proc. of IMTC'99 Conference*, 1999, pp. 880-885.
- [40] J.L González, Eduard Alarcón, "Study of Delta-I Noise Effects on the Spectral Performance of Current-Steering Segmented Digital-to-Analog Converters", *Proceedings of the 5th IEEE Workshop on Signal Propagation on Interconnects (SPI'01)*, Cavallino-Venice, Italy, May 2001.
- [41] A. Van den Bosch, M. Steyaert, W. Sansen, "Design Techniques for High Accuracy, Current-Steering CMOS D/A Converters", *3rd Workshop on ADC Modelling and Testing*, 1998, pp. 861-866.
- [42] Hiroyuki Kohno, Yasuyuki Nakamura, Atsuhito Kondo, Hiroyuki Amishiro, Takahiro Miki, Keisuke Okada, "A 350-MHz 8-bit CMOS D/A Converter Using Delayed Driving Scheme", *Proc. of IEEE 1995 Custom Integrated Circuits Conference*, pp. 211-214, 1995.
- [43] D. Johns, K. Martin, *Analog Integrated Circuit Design*, Wiley & Sons, 1997.
- [44] Seung-Chul Lee, Jin Park, Jin-Sik Yoon, Jung-Hee Song, and Seung-Hoon Lee, "A 3 V 10b 100 MS/s Digital-to-Analog Converter for Cable Modem Applications", *Proc. of IEEE 1998 ICECS*, 1998, pp. 203-205.

Chapter 4

CMOS Comparators

Rafael Domínguez-Castro, Manuel Delgado-Restituto, Angel Rodriguez-Vázquez, José M. de la Rosa, and Fernando Medeiro

4.1 INTRODUCTION^{†1}

Comparators are used to *detect* whether an *analog* signal $x_+(t)$ is larger or smaller than other $x_-(t)$ ^{†2}, and to *codify* the outcome in digital domain as follows,

$$y = \begin{cases} \mathbf{1}_D & \text{for } x_+(t) > x_-(t) \\ \mathbf{0}_D & \text{for } x_+(t) < x_-(t) \end{cases} \quad (4.1)$$

where y is the output signal, $\mathbf{0}_D$ represents the logic zero and $\mathbf{1}_D$ is the logic one. Ideal comparators should be capable to detect arbitrarily small differences between the input signals. However, in practice, these differences must be larger than a characteristic *resolution* parameter ξ for proper detection. For a given comparator circuit, the value of this resolution parameter changes depending upon the operating conditions. If the temporal window allocated for comparison is long enough, ξ takes an absolute minimum value which is inherent in the comparator device and which defines its maximum accuracy. As the temporal window shrinks, the value of ξ increases above its absolute minimum value and, hence, the comparator accuracy worsens. It highlights a trade-off between *accuracy* and

-
1. **Acknowledgment:** This work has been partially supported by the spanish MCyT and the ERDF – Project TIC2001-0929 (ADAVERE), and the European Union Project IST-2001-34283 (TAMES). The useful comments from Dr. Gustavo Liñán are highly appreciated.
 2. In many applications one of the inputs is a reference value, say $x(t) = E$, and the comparator detects whether the signal applied to the other input, say $x_+(t) \equiv x(t)$, is larger or smaller than such reference.

speed of operation; the larger the speed the smaller the accuracy. As in any other analog circuit this trade-off is also influenced by power consumption and area occupation.

Comparators are basic building blocks of analog-to-digital converters. They are hence crucial components to realize the front-ends of the newest generations of *mixed-signal* CMOS electronic systems [1] [2]. Other comparator applications include such diverse areas as signal and function generation [3], digital communications [4], or artificial neural networks [5], among others.

This chapter first presents an overview of CMOS voltage^{†3} comparator architectures and circuits. Starting from the identification of the comparator behavior, Section 2 introduces several comparator architectures and circuits. Then, Section 3 assumes these topologies, characterizes high-level attributes, such as static gain, unitary time constant, etc., and analyzes the resolution – speed trade-off for each architecture. Such analysis provides a basis for comparison among architectures. These previous sections of the chapter neglect the influence of circuit dissymmetries. Dissymmetries are covered in Section 4; and new comparator topologies are presented to overcome the offset caused by dissymmetries. Related high-level trade-offs for these topologies are also studied in this section.

4.2 OVERVIEW OF BASIC CMOS VOLTAGE COMPARATOR ARCHITECTURES

Fig. 4.1(a) shows the static transfer characteristic of an ideal comparator where $x \equiv x_+(t) - x_-(t)$ and E_{OH} and $-E_{OL}$ are levels that correspond to the logic one and zero, respectively. From now on, we will implicitly assume that comparator inputs and output are voltages. The case where inputs are currents and the output

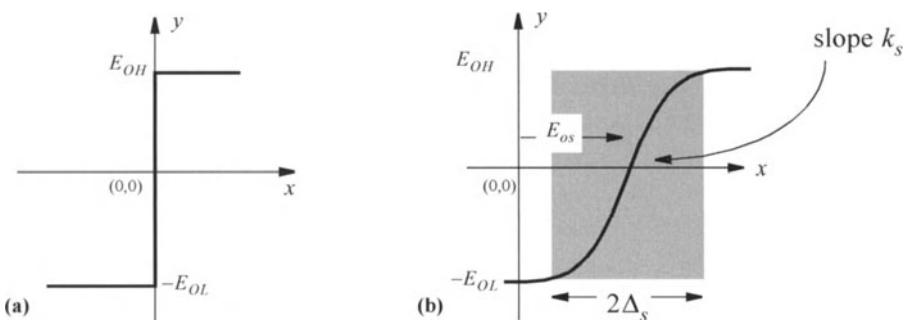


Figure 4.1: Comparator transfer characteristics: (a) Ideal; (b) First-order non-

3. This is used here to mean that inputs and output are all voltages.

is a voltage – current comparators – will also be considered in this chapter.

From Fig. 4.1(a) it is seen that an ideal voltage comparator must exhibit infinite *voltage gain* around the zero value of the differential input x . Obviously, this cannot be achieved by any real device. Fig. 4.1(b) shows a better approximation of the static transfer characteristic exhibited by actual comparators. There, the transfer characteristic is assumed to have a *finite* static gain, k_s , around an *input offset* voltage, E_{os} . On the basis of this nonlinear characteristic, the minimum value of the resolution, called herein *static resolution*, is,

$$\xi_s \approx |E_{os}| + \frac{E_{OH} + E_{OL}}{2k_s} \approx |E_{os}| + \left. \frac{E_{OH}}{k_s} \right|_{\text{for } E_{OL} \approx E_{OH}} \quad (4.2)$$

where the random nature of the offset has been accounted for by including its module, and it has been assumed that the shaded transition interval is symmetrical around the input offset ^{†4}. For any input level inside the interval $[-\xi_s, \xi_s]$ the comparator digital output state is uncertain. Otherwise, any input level outside this interval, called an *overdrive*, generates an unambiguous digital state at the comparator output. The overdrive variable measures how far from this interval the actual input is, $|x_{ovd}| = |x| - \xi_s$.

4.2.1 Single-Step Voltage Comparators

Voltage comparators are basically *voltage gain* devices. Hence, they can be implemented with the same circuit topologies employed for voltage amplifiers. Figs. 4.2(a)-(c) shows three CMOS alternatives which are all *single-stage Operational Transconductance Amplifiers* (OTA) [6]. Those in Figs. 4.2(a) and (b) have differential input and single-ended output and will be called, respectively, AOTAC (Asymmetric Operational Transconductance Amplifier Comparator) and SOTAC (Symmetric Operational Transconductance Amplifier Comparator). On the other hand, that in Fig. 4.2(c) is a fully-differential topology where the output is obtained as the difference between the voltages at the output terminals of a symmetrically loaded differential pair. In this structure, called FDOTAC (Fully-Differential Operational Transconductance Amplifier Comparator), the bias current of the differential-pair must be controlled through a feedback circuitry in order to stabilize and set the quiescent value of the common-mode output voltage

4. For more accurate calculation of ξ_s , the levels which guarantee unambiguous interpretation of the logic zero and one, namely $E_{RH} < E_{OH}$ and $-E_{RL} > -E_{OL}$, should be used instead of E_{OH} and $-E_{OL}$. Also, the gain should not be considered constant over the transition interval. Finally, two different transition intervals should be considered, one for positive excursions Δ_{s+} and another for negative excursions Δ_{s-} .

[7]; this common-mode regulation circuitry has not been included in Fig. 4.2(c).

All circuits in Figs. 4.2(a)-(c) use the same mechanism for achieving voltage gain. Fig. 4.2(d) shows a first-order conceptual model for approximating the behavior underlying such mechanism around the input transition point. There, the transconductance g_m models the operation of the differential pair formed by the matched^{†5} transistors M_{Nd} and the associated biasing transistors M_{NB} , the resistance r_o models the combined action of the output resistances of all transistors, and the capacitance C_o models the combined effect of external capacitive loads and transistor parasitics at the output node. The three first rows in Table 4.1: include expressions for these model parameters in terms of the transistor sizes, the

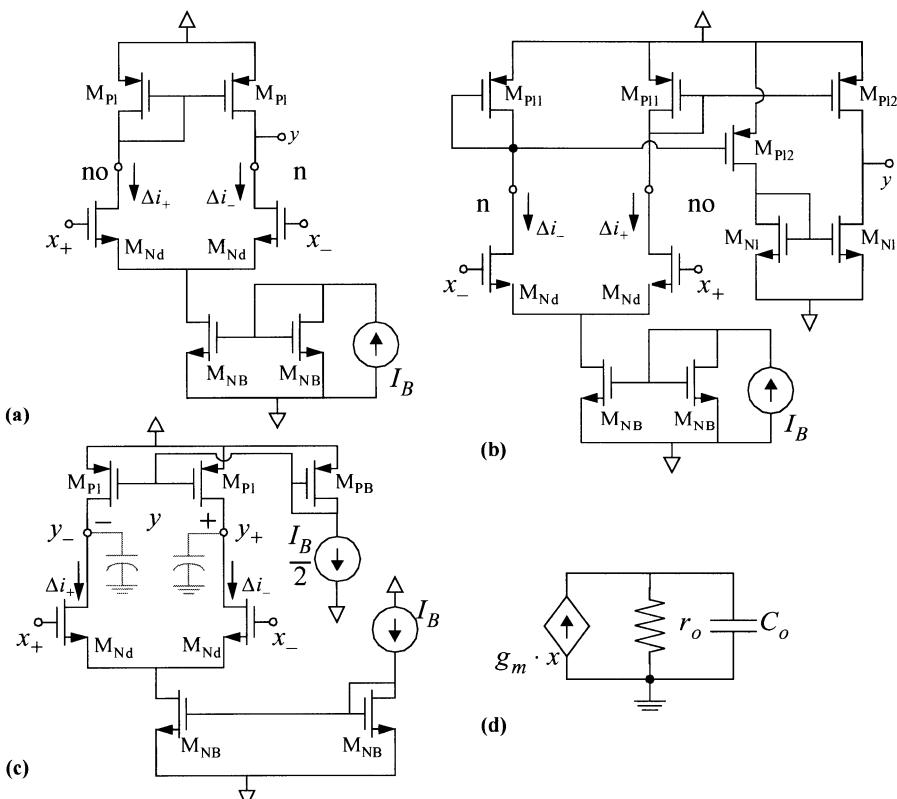


Figure 4.2: (a)-(c) CMOS Single-step comparator topologies; (d) First-order small-signal model around the input offset voltage.

-
5. The term matched here means that the transistors are designed to be equal; i.e. same sizes, same orientation, same surrounding, etc. In practice, the transistors are mismatched. Actually, mismatch is the main source of input offset.

large-signal MOST transconductance density β_0 , and the MOST Early voltage V_A – see Appendix 6 for a simplified MOST model.

In all three circuits of Fig. 4.2 the static gain is given by,

$$k_s = g_m \cdot r_o \quad (4.3)$$

which, using the expressions of Table 4.1:, and assuming that Early voltages are proportional to transistor lengths, yields the following dependence of the static gain on design parameters,

$$k_s \propto \frac{\sqrt{W_{Nd}} \cdot \sqrt{L}}{\sqrt{I_B}} \quad (4.4)$$

These structures are well suited to provide static gains around 40dB, which from (4.2), neglecting for the moment the input offset, and assuming $E_{OH} \approx 1\text{V}$, results in an absolute value of the static resolution of around 10mV.

The limited resolution due to low static gain values can be overcome by resort-

Table 4.1: Model parameters of one step CMOS comparator structures.

Param. → Structure ↓	g_m	$g_o = r_o^{-1}$	C_o
AOTAC		$\frac{I_B}{2} \cdot \left(\frac{1}{V_{ANd}} + \frac{1}{V_{APl}} \right)$	$C_L + v_{Nd} \cdot W_{Nd} \cdot C_{GD0Nd} + v_{Pl} \cdot W_{Pl} \cdot C_{GD0Pl}$
SOTAC		$\frac{I_B}{2} \cdot \left(\frac{1}{V_{ANI}} + \frac{1}{V_{APl2}} \right)$	$C_L + v_{NI} \cdot W_{NI} \cdot C_{GD0NI} + v_{Pl2} \cdot W_{Pl2} \cdot C_{GD0Pl2}$
FDOTA C	$\sqrt{2 \cdot \frac{\beta_{0Nd}}{n_{Nd}} \cdot \left(\frac{W}{L} \right)_{Nd} \cdot I_B}$	$\frac{I_B}{2} \cdot \left(\frac{1}{V_{ANd}} + \frac{1}{V_{APl}} \right)$	$C_L + v_{Nd} \cdot W_{Nd} \cdot C_{GD0Nd} + v_{Pl} \cdot W_{Pl} \cdot C_{GD0Pl}$
FOTAC		$\frac{1}{1 \cdot V_{ANI} \cdot V_{ACN}} \sqrt{\frac{I_B^3}{2 \cdot \frac{\beta_{0CN}}{n_{CN}} \cdot \left(\frac{W}{L} \right)_{CN}}} + \frac{1}{4 \cdot V_{APl} \cdot V_{ACP}} \sqrt{\frac{I_B^3}{2 \cdot \frac{\beta_{0CP}}{n_{CP}} \cdot \left(\frac{W}{L} \right)_{CP}}}$	$C_L + v_{CN} \cdot W_{CN} \cdot C_{GD0CN} + v_{CP} \cdot W_{CP} \cdot C_{GD0CP}$

ing to the use of cascode transistors. The circuit of Fig. 4.3(a), labelled FOTAC (Folded Operational Transconductance Amplifier Comparator), is a representative example. The first-order model of Fig. 4.2(d) is still valid but now the resistance parameter r_o is increased in a factor approximately equal to the gain of the cascode devices M_{CN} and M_{CP} . Then, assuming that all transistors have the same channel length, it follows that,

$$k_s \propto \frac{\sqrt{W_{Nd}} \cdot \sqrt{W_C} \cdot \sqrt{L^3}}{I_B} \quad (4.5)$$

which renders this structure appropriate for obtaining static gains up to around 80dB and, hence, Δ_s around 0.1mV. Further gain enhancement can be achieved by enforcing the cascode action through the incorporation of local feedback amplifiers – illustrated in Fig. 4.3(b) [8].

4.2.2 Multi-Step Comparators

Essentially, in the structures of Figs. 4.2 and 3, the gain needed for the comparison function is built in a *single step*^{†6}. Later in the chapter it will be shown that this results into a disadvantageous resolution – speed trade-off. To relax this trade-off, multi-step structures are employed that achieve the voltage gain into multiple steps, through the multiplication of several $g_m \cdot r_o$ gain factors [9] [10]. Fig. 4.4(a) shows such a multi-step architecture at the conceptual level. Assuming for illustration that the N stages are identical, each having a gain $k_s = g_m / g_o$, the following expression is calculated for the static resolution,

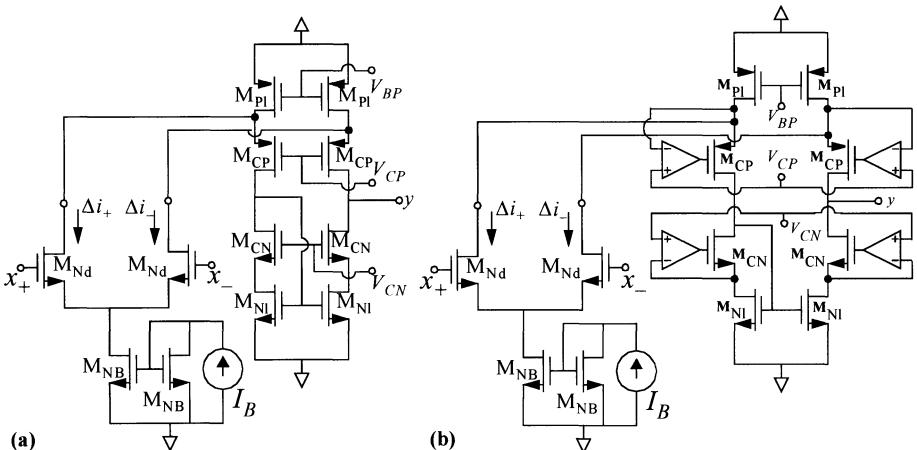


Figure 4.3: Single-step comparators with cascode transistors.

$$\xi_s \approx |E_{os}| + E_{OH} \cdot \left(\frac{g_o}{g_m} \right)^N \quad (4.6)$$

where $|E_{os}|$ is the offset voltage of the first stage of the cascade. The offset of each of the remaining stages is neglected because it is attenuated by the gain of the preceding stages in the chain. In comparison with (4.2), (4.6) shows that the impact of the static gain on the static resolution is much smaller than for single-step architectures. Actually, for large enough N the static resolution of multi-step

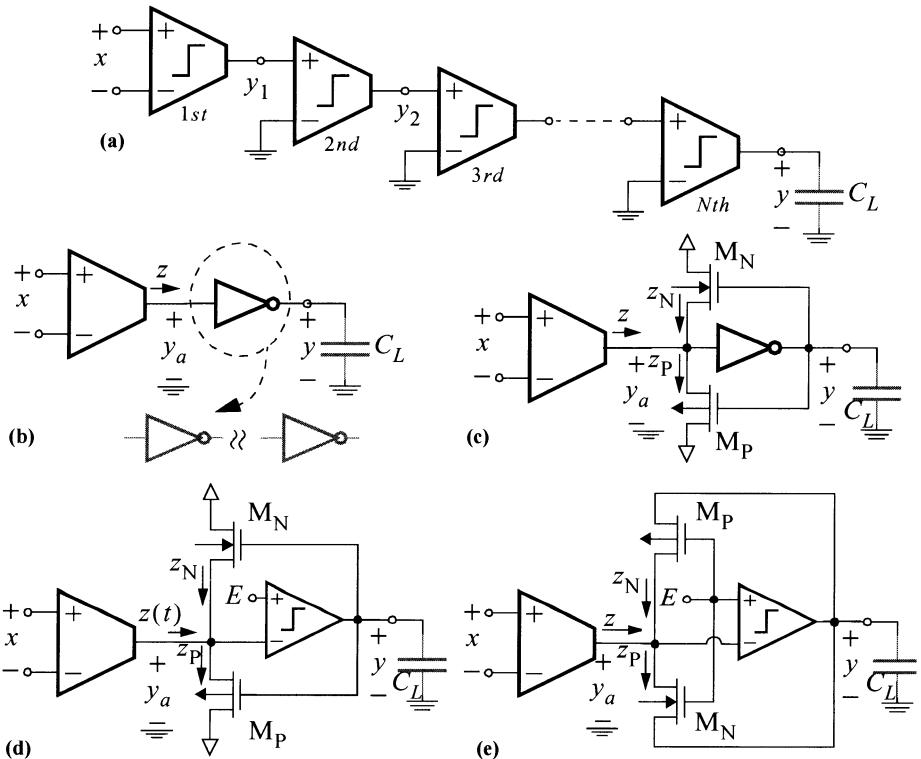


Figure 4.4: Multi-step voltage comparators at the conceptual level.

6. We understand that the basic mechanism to achieve voltage gain is multiplying a small-signal transconductance g_m by a small-signal resistance r_o . The product $g_m \cdot r_o$ defines the basic voltage gain factor. In the simplest OTA the gain is equal to the product of just one transconductance by one resistance; i.e. it is equal to single gain factor $g_m \cdot r_o$. This is the reason why we say that these structures obtain the gain into one step. It can be argued that this does not apply to cascode architectures, where the gain is enhanced through multiplying r_o by the gain of the cascode transistors. However, for convenience we also consider cascode architectures as single-step comparators.

architectures is basically limited by the offset voltage, and the influence of the static gain becomes negligible.

The stages employed in a multi-step comparator are generically different. Fig. 4.4(b) shows a typical topology consisting of a front-end OTA followed by a CMOS inverter. Similar to what is done for buffering logic signals, several CMOS inverters with properly scaled transistor dimensions can be cascaded to enhance the speed of the logic transitions for given capacitive load – as illustrated in the figure inset. Fig. 4.5(a) shows an actual CMOS implementation of one of such topologies [11].

We can think of a CMOS inverter as a current comparator [12]. Any positive

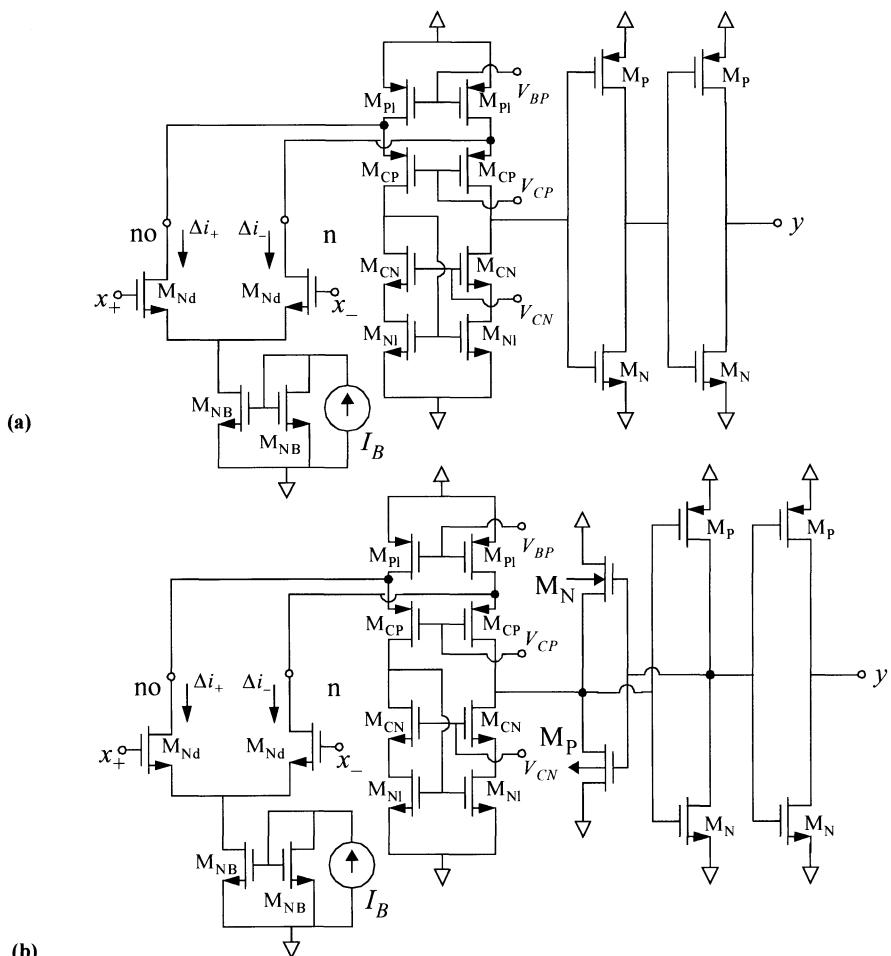


Figure 4.5: Some multi-step CMOS voltage comparators.

current driving the inverter is integrated by the input capacitor – see Fig. 4.4(b) – increasing voltage y_a , and driving the output voltage to the logic zero; reciprocally, any negative current driving this node makes voltage y_a decrease, and drives the output to the logic one. Thus, the operation of the two-step topology of Fig. 4.4(b) can be described as follows: the front-end OTA transforms the input voltage difference x into a current $z = g_m \cdot x$, whose sign is then detected by the CMOS inverter operating as a current comparator. Based on this view, improved multi-step comparator architectures can be obtained by simply replacing the CMOS inverter by high-performance CMOS current comparators [12].

Fig. 4.4(c) presents a circuit based on this principle that requires just two extra transistors, labelled M_N and M_P . This circuit has three operating regions. For small changes around the *quiescent point*^{†7} ($z = 0$ and $y_a \approx 0$), transistors M_N and M_P are OFF, the inverter practically operates in open-loop configuration, and node y_a is capacitive. For $z > 0$, the voltage y_a is pulled up and the amplifier forces the voltage at node y to decrease; thus, M_P becomes ON, and the subsequent feedback action clamps the voltage y_a to,

$$y_a \approx \frac{|V_{TP}|}{1 + A_0} \equiv \Delta_H \quad (4.7)$$

where A_0 denotes the inverter gain, and V_{TP} stands for the threshold voltage of the PMOS device. The dual situation occurs for $z < 0$, when y_a is pulled down, y increases, M_N becomes ON, and the feedback action clamps y_a to,

$$y_a \approx -\frac{V_{TN}}{1 + A_0} \equiv -\Delta_L \quad (4.8)$$

Thus, it is seen that y_a changes into a rather small interval $[-\Delta_L, \Delta_H]$, which may render important speed advantages. Fig. 4.4(d) achieves better control of the location of this interval, together with further reduction of its amplitude, by replacing the inverter by a differential input amplifier. There, the virtual ground action of this amplifier forces y_a to remain very close to E , regardless of the magnitude of the input signal difference. A proper choice of the value of E can help to improve the operation of the front-end OTA by conveniently biasing its output transistors.

The structures of Figs. 4.4(c) and (d) have the drawback that the transient behavior is largely dominated by the capacitor across the nodes y_a and y [11]. The circuit of Fig. 4.4(e) circumvents this problem by decoupling the amplifier

7. Quiescent point is used across this paper to denote the point around which the comparator gain is built. We assume that comparators are perfectly balanced when biased at the quiescent point, and that detection happens around this point.

input and output nodes. Its static operation follows similar principles to those of Fig. 4.4(d). When $z = 0$, transistors M_P and M_N are OFF and the circuit exhibits capacitive-input behavior. Positive currents are integrated in the input capacitor increasing voltage y_a and, consequently, decreasing y until the transistor M_N becomes conductive, drawing the input current and stabilizing the output. The other way around, M_P is the conductive transistor for negative input currents.

A common feature of the comparator circuits in Figs. 4.4(c)-(e) is that the output voltage swings a rather limited interval of amplitude $\approx V_{TN} + |V_{TP}|$. Then, additional CMOS inverters in series with the output node may be necessary in order to restore the logic levels, as shown in Fig. 4.5(b).

4.2.3 Regenerative Positive-Feedback Comparators

Although the very operation of voltage comparators consists of building voltage gain stages, there are significant differences among amplifiers themselves and comparators. Amplifiers are usually employed to achieve linear operation in closed-loop configurations, which requires careful compensation of the dynamic response to avoid unstable operation when feedback is applied. On the contrary, the dynamic of the gain mechanism employed for comparators does not even need to be stable in open loop. Actually, *positive feedback* can be employed to implement mechanisms with *unstable*^{†8}, very fast gain-building mechanisms [13]. Actually, in the next section it will be shown that regenerative comparators are inherently faster than other types.

Consider, for illustration purposes, the incorporation of positive feedback to single-step comparators. Fig. 4.6(a) shows a circuit implementation where positive feedback action is exercised by the OTA, whose small-signal transconductance is g_{mpf} , during the *active clock phase* φ_a . At this point, let us accept without explanation the use of clock-controlled switches in this circuit. Note that the controlling clock has two non-overlapped phases, as shown in Fig. 4.6(a). Comparisons take place only when the clock phase φ_a is in the high state, and consequently switches controlled by this phase are ON, while the others are OFF.

Fig. 4.6(b) shows a first-order model to represent the behavior of Fig. 4.6(a) around its quiescent point $y = 0$, when the clock phase φ_a is in the high state. Positive feedback is modeled by the negative resistance $r_{pf} = -1/g_{mpf}$ which counterbalances the negative feedback action exercised by the resistance r_o . Provided that $g_{mpf} > g_o = 1/r_o$, the global feedback around the quiescent point is positive and, hence, the global behavior is unstable.

8. Here unstable means that the small-signal model around the quiescent point has poles in the right-hand side of the complex frequency plane.

For better understanding of the qualitative operation of Fig. 4.6(a), it is helpful to consider the plots of Figs. 4.6(c) and 6(d). There, we depict the approximate resistive characteristics “seen” by capacitor C_o during the comparison phase due to the combined action of the two OTAs of Fig. 4.6(a), and taking into account the OTA non-linearities ^{†9}. Fig. 4.6(c) corresponds to “small” values of the input x , whereas Fig. 4.6(d) corresponds to “large” input values. The actual characteristic seen in each case depends on the sign of the input applied during the comparison phase; continuous traces correspond to positive input trajectories whereas dashed traces represent negative input trajectories. Independently of the input being large or small, positive or negative, during the *reset* phase the output is driven to the central point P_0 , for which $y = 0$. This defines the quiescent point, where the small-signal model of Fig. 4.6(b) is applicable.

Consider now that a “small” input is applied during the comparison phase. For $x > 0$, the capacitor sees the bottom characteristics of Fig. 4.6(c) which includes three *equilibrium points* ^{†10}: two stable, Q_L and Q_H , and the other unstable, Q_0 . Because the capacitor charge cannot change instantaneously, the initial state at $y = 0$ corresponds to the point P_+ on the characteristic, which is located on the right-hand side of Q_0 . From P_+ the repulsion action exercised by Q_0 , precludes the left-hand stable equilibrium at Q_L to be reached, and the trajectory becomes

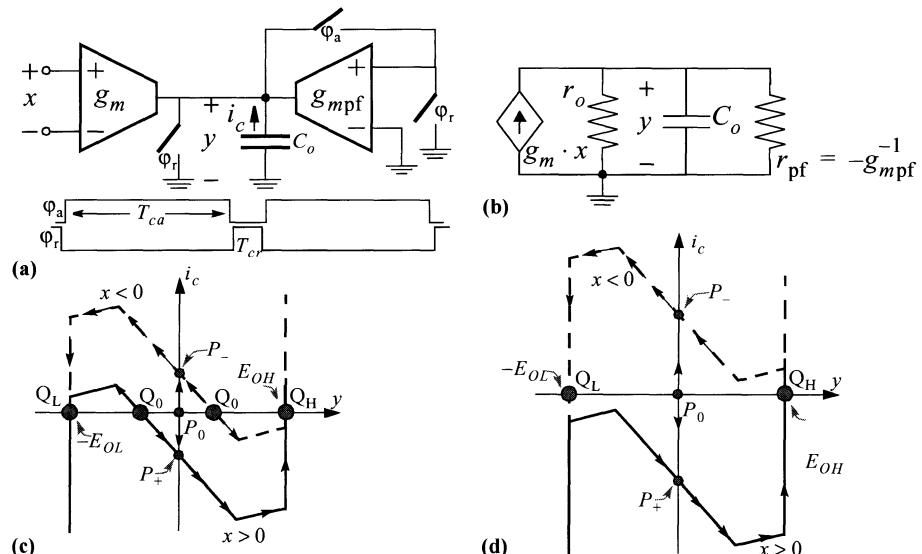


Figure 4.6: Including positive feedback into single-step comparators.

9. Here two basic non-linearities are considered, namely saturation of the OTA transconductance and saturation of the OTA output voltage [14].

attracted towards the right-hand stable equilibrium at Q_H , where $y = E_{OH}$. On the other hand, for $x < 0$, the central point pushes the trajectory towards the equilibrium at Q_L , where $y = -E_{OL}$. In both cases, the dynamic evolution around the central point is governed by the model of Fig. 4.6(b) and, hence, realized at high-speed due to the positive feedback action.

For “large” input values, the characteristic of Fig. 4.6(d) applies. In such a case, there is only one stable equilibrium point for each sign of x , and the description of the transient evolution is similar to the previous one.

Note that the qualitative description above remains valid for whatever input magnitude. It suggests that, unlike for single-step topologies, the resolution of this type of comparators is not limited by the static gain. Unfortunately, the offset limitation is more important here [14].

At this point, we have the ingredients needed to understand why the circuit of Fig. 4.6(a) is clocked: if the circuit were operated in continuous-time instead of discrete-time, *hysteresis* would arise. In order to understand this, let us return to Figs. 4.6(c) and (d). Consider that at a certain time instant, a positive input is applied such that the circuit is at the stable equilibrium point Q_H , and consequently the output is high. Assume now that the input decreases. Fig. 4.6(c) tells us that Q_H remains as a stable equilibrium point even when the input becomes negative; only for large enough negative inputs, Q_H ceases to be a stable equilibrium point, as Fig. 4.6(d) illustrates. It means that large enough negative values must be applied to counterbalance the circuit “inertia” to remain in the high-state and, thereby, force its evolution towards the low state. This inertia, which is a consequence of the circuit memory, is eliminated by employing switches to place the circuit at the unstable equilibrium point P_0 before comparison is actually made [14]. This operation, realized during the reset clock phase, is equivalent to erasing the memory of the circuit and it is the key to guarantee that hysteresis will not appear.

Discrete-Time (DT) regenerative comparators are commonly built by cross-coupling a pair of inverters to form a *latch* – a circuit structure often used as sense amplifier in dynamic RAMs [16]. Fig. 4.7(a) shows the concept of regenerative comparison based on a latch, where the blocks labelled “ τ ” model delays in the

10. At the intersection points the current through the capacitor is null and hence $dy/dt = 0$.

These points are equilibrium states where $y(t) = \text{cte}$ and the circuit may remain static [15]. In practice, the circuit will actually remain static as long as the slope of the i_o vs y curve is positive around the point (stable equilibrium) and will not otherwise (unstable equilibrium). Starting from any arbitrary initial value of y , the circuit trajectory towards steady-state is determined by the attraction exercised by stable equilibrium points, and the repulsion exercised by unstable equilibrium points.

transmission of voltages around the feedback loop. The inverters amplify the differential input $x_{a+} - x_{a-}$ to obtain the saturated differential output $y_+ - y_-$ according to the characteristics drawn with solid line in Fig. 4.7(b). During the reset phase (φ_r high), the differential input is stored at the input sampling capacitors and the circuit is driven to the central state Q_0 . During the active phase (φ_a high), the differential input is retrieved, forcing an initial state either on the right ($x > 0$) or on the left ($x < 0$) of Q_0 . From this initial state, the action of positive feedback forces the output to evolve either towards Q_H , for $x > 0$, or towards Q_L , for $x < 0$, as illustrated through the grey-line trajectories in Fig. 4.7(b).

Figs. 4.8(a) to (d) show some examples of CMOS latches reported in literature [17] [18] [19] [20] [21]. For those in Figs. 4.8(a) and (b), transistors M_{NB} and

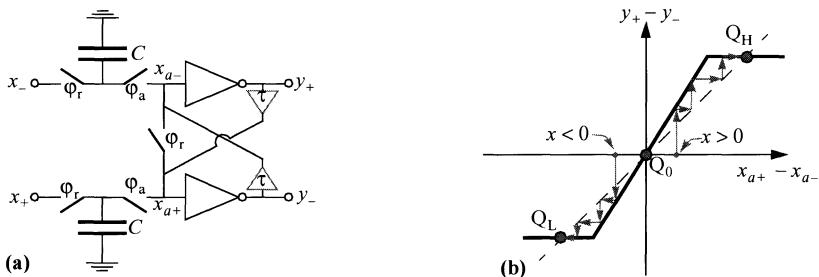


Figure 4.7: Conceptual latch circuit model and illustrative dynamic operation.

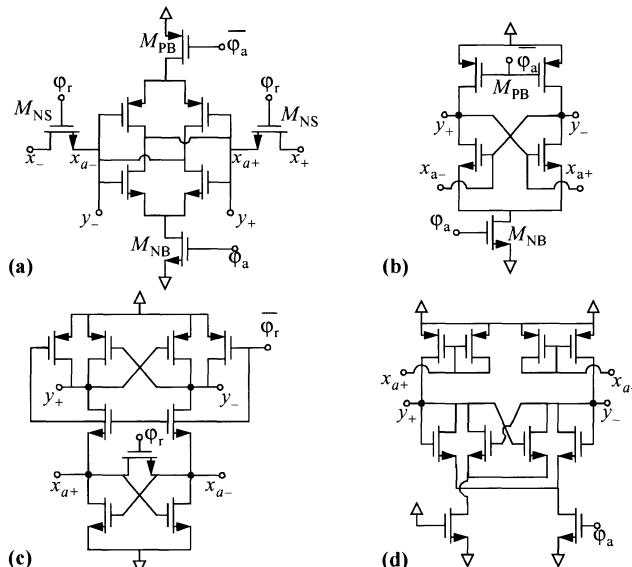


Figure 4.8: Some CMOS latches.

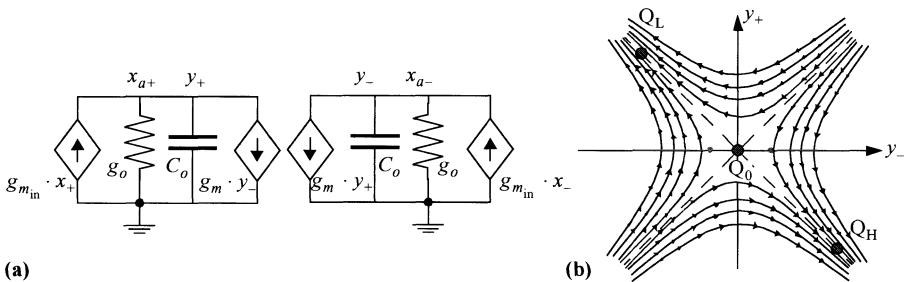


Figure 4.9: (a) Small-signal model around the quiescent point. (b) State diagram

M_{PB} are OFF during the reset phase, so that the latch is disabled. In this phase, nodes x_{a+} and x_{a-} are in high-impedance state and input voltages can be sampled therein; transistors M_{NS} in Fig. 4.8(a) are used for that purpose. Next, the voltage difference is amplified when the latch is enabled during the active phase. Alternatively, nodes x_{a+} and x_{a-} can be driven during the active phase with currents obtained from the input voltages by means of transconductors. This is the unique excitation alternative for the latches of Figs. 4.8(c) and (d).

Fig. 4.9(a) shows a first-order small-signal model for the CMOS latches of Figs. 4.8(c) and (d) around the quiescent point. The action of the excitation transconductors mentioned above, represented by transconductances $g_{m_{in}}$, is also modeled for completeness, although they are not shown in the actual CMOS circuits of Fig. 4.8. Fig. 4.9(b) shows the state diagram (y_+ vs. y_-) with the three equilibrium points of the system ^{†11} and different dynamic trajectories around the quiescent point. It is seen that these trajectories are “separated” by the bisecting lines so that half of the plane converges towards Q_H and the other half towards Q_L . Hence, following the setting of the system at Q_0 during reset phase, the unbalance established during comparison phase combined with the separating action exercised by the unstable equilibrium point will force a transient evolution towards the correct stable equilibrium point.

4.2.4 Pre-Amplified Regenerative Comparators

Ideally, the static resolution of regenerative comparators is unlimited; i.e., even arbitrarily small input unbalances could be detected ^{†12}. In practice, their resolu-

11. The small-signal model only accounts for the central, unstable equilibrium point Q_0 . Nonlinearities must be brought into the picture in order to account for the stable equilibrium points Q_L and Q_H .

12. Even in such an ideal case, detection of infinitely small inputs would require infinitely long detection times, as the formulae of the next section will show.

tion is limited by dissymmetries and other second-order phenomena. Moreover, errors caused by dissymmetries are much larger in comparators with positive feedback than in other types [1] [9]. Thus, in order to keep the speed advantages of regenerative comparator and simultaneously improving its resolution, a pre-amplifier is usually placed in front of the regenerative core.

This is the strategy employed in the conceptual schematic of Fig. 4.10(a). There, the inverters in the latch are self-biased during the reset phase. During the active phase, the input signal $x_+ - x_-$ is first amplified by a factor A_{in} , and then added to the quiescent input voltages of the inverters prior to enable the positive feedback loop. Input signal amplification occurs during the first part of the active phase (φ_{a1} high). Positive feedback is enabled during the last portion of the active phase (φ_{a2} high), whose rising edge is delayed with respect to that of φ_{a1} in order to guarantee that a *large enough* input unbalance is built prior to closing the positive feedback loop. Thus, the resolution improvement of this architecture as compared to the one without pre-amplifier is roughly proportional to A_{in} .

Preamplification is also the role played by the two transconductances labelled $g_{m_{in}}$ in the model of Fig. 4.9(a). Fig. 4.10(b) shows a CMOS circuit implementation of this concept using the CMOS latch of Fig. 4.8(c) [17].

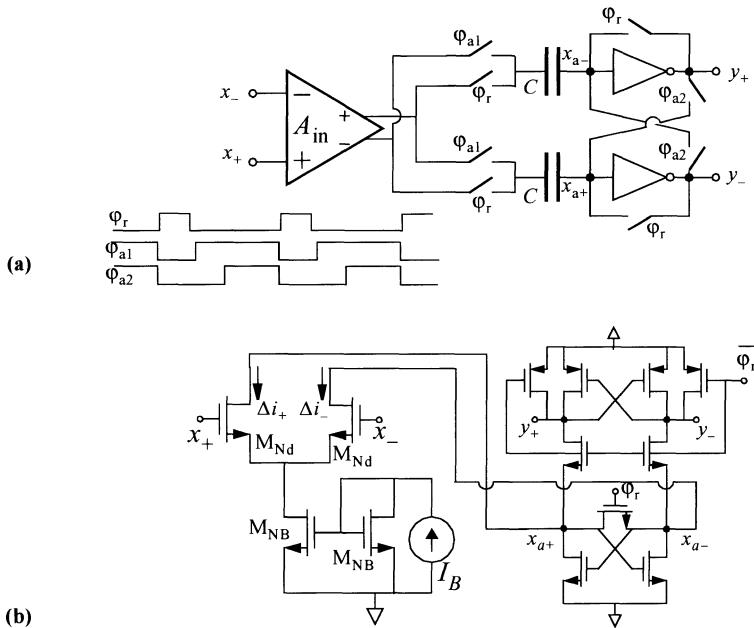


Figure 4.10: Pre-amplified regenerative comparators.

4.3 ARCHITECTURAL SPEED-VS-RESOLUTION TRADE-OFFS

The resolution – speed trade-off can be studied at different levels. At the transistor level, we would like to know how speed and resolution are influenced by design parameters such as transistor sizes and bias current. Instead, this section follows another approach where comparators are characterized by high-level parameters, such as static gain, unitary time constant, etc., and the trade-off is formulated as a function of these parameters. This approach enables us to draw comparisons among the architectures.

For any of the comparator topologies mentioned so far, for instance Fig. 4.4, let us assume that the comparator is perfectly balanced for $x = 0$; i.e., that the output is in the quiescent point with $y = 0$. It means that input offsets are ignored by now. Let us then consider that a step stimulus of value Δ_d is applied at $t = 0$. By using small-signal models around the quiescent point, the output waveform can be calculated as,

$$y(t) = \Delta_d \cdot k_d(t) \quad (4.9)$$

where $k_d(t)$, called *dynamic gain*, depends upon the actual comparator structure being considered. For each structure, $k_d(t)$ reflects the time needed to build a given amount of gain. Particularly, for a given Δ_d we are interested in knowing the time needed to build the necessary gain so that the output reaches the restoring logic level. This time, called *quiescent comparison time* T_c , is calculated from,

$$E_{OH} = \Delta_d \cdot k_d(T_c) \quad (4.10)$$

This equation defines a resolution – speed trade-off which is different for each architecture. Generally speaking, the larger T_c – i.e., the slower the comparator the smaller Δ_d – i.e., the more accurate the comparator. However, the actual dependence between T_c and Δ_d changes from one architecture to another, as the curves depicted in Fig. 4.11 suggest. The meaning of these curves is explained in the following sections.

4.3.1 Single-Step Comparators

Let us first focus on single-step voltage comparators, and assume that the dynamic behavior around the quiescent point is represented by the model of Fig. 4.2(d). This first-order model accounts for the static gain and the unitary frequency of the gain mechanism. Consider that the capacitor C_o in Fig. 4.2(d) is discharged at $t = 0$ and that a step excitation of amplitude Δ_d is applied around the input threshold voltage. The output waveform is,

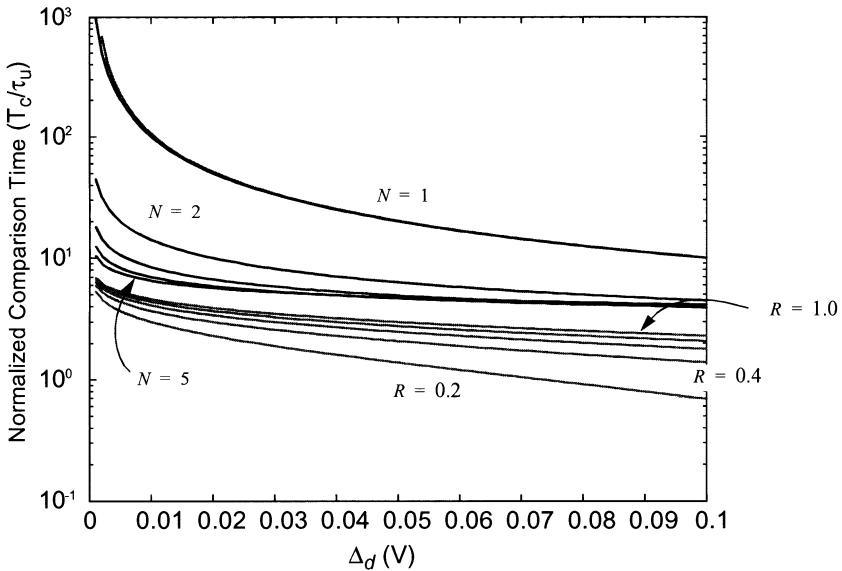


Figure 4.11: Illustrating the resolution-speed trade-off.

$$y(t) = \Delta_d \cdot k_s \cdot \left(1 - e^{-\frac{t}{\tau_o}} \right) , \quad \tau_o \equiv \frac{C_o}{g_o} \quad (4.11)$$

Assume now that Δ_d is large enough in comparison with Δ_s defined in Fig. 4.1(b), so that $\Delta_d \cdot k_s \gg E_{OH}$. Then, analysis of (4.11) shows that $y(t)$ reaches the restoring level E_{OH} into a small fraction of τ_o . Taking this into account, (4.11) can be series-expanded and approximated to obtain,

$$y(t) \approx \Delta_d \cdot k_s \cdot \frac{t}{\tau_o} \quad (4.12)$$

Comparing this result with (4.9), the dynamic gain is expressed as,

$$k_d(t) = k_s \cdot \frac{t}{\tau_o} = \frac{g_m}{g_o} \cdot \frac{g_o}{C_o} \cdot t \equiv \frac{t}{\tau_u} \quad (4.13)$$

where $\tau_u \equiv C_o/g_m$ is the *unitary time constant* of the amplifier. Correspondingly, the resolution – speed trade-off is given by,

$$\Delta_d \cdot \frac{T_c}{\tau_u} \approx E_{OH} \quad (4.14)$$

The curve labelled as $N = 1$ in Fig. 4.11 illustrates this trade-off for $E_{OH} = 1\text{V}$. As Δ_d decreases, (4.14) shows that T_c increases at the same rate. However, this equation is valid only if $\Delta_d \cdot k_s \gg E_{OH}$. As the input approaches the static resolution limit, i.e. as $\Delta_d \rightarrow \Delta_s \equiv E_{OH}/k_s$, it is not possible to assume $T_c \ll \tau_o$, and the trade-off is recalculated as,

$$\Delta_d \cdot \frac{T_c}{\tau_u} = E_{OH} \cdot \left[\left(\frac{k_s \cdot \Delta_d}{E_{OH}} \right) \ln \frac{1}{1 - \frac{1}{k_s} \cdot \frac{E_{OH}}{\Delta_d}} \right] \quad (4.15)$$

Consider $\Delta_d \approx \Delta_s \cdot (1 + \varepsilon)$ with $\varepsilon \ll 1$. (4.15) can be simplified to obtain a relationship between the static gain and the time needed to obtain such limiting sensitivity,

$$\frac{T_c}{\tau_u} = A_0 \cdot \ln \left(\frac{1}{\varepsilon} \right) \quad (4.16)$$

where for the sake of homogeneity with the text in the following subsection, the static gain has been renamed as A_0 . In the limit, as $\Delta_d \rightarrow \Delta_s$ and $\varepsilon \rightarrow 0$, it follows that $T_c \rightarrow \infty$.

Let us consider for illustration purposes an example with $k_s = 2 \times 10^3$, $\tau_u = 10\text{ns}$ and $E_{OH} = 1\text{V}$. Thus, $\Delta_d = 10\text{mV}$ requires from (4.14) $T_c \approx 1\mu\text{s}$, while $\Delta_d = 1\text{mV}$ requires from (4.15) $T_c \approx 14\mu\text{s}$. On the other hand, if the static resolution limit has to be approached within 1%, (4.16) yields $T_c \approx 92\mu\text{s}$.

Summing up, we can state that the single-step comparator, under normal operation conditions (i.e. $\Delta_d \ll E_{OH}$), reacts at much lower speed than the underlying gain mechanism; i.e., T_c is much larger than τ_u .

4.3.2 Multi-Step Comparators

Let us focus now on multi-step comparators – see Fig. 4.4. Consider for illustration purposes that all stages are identical, ignore again the input offset voltages, and assume that all capacitors are discharged at $t = 0$, and that an input step valued Δ_d is applied at that time instant. Laplace-domain analysis shows that,

$$Y(s) = \left(\frac{k_s}{1 + s \cdot \tau_o} \right)^N \cdot \frac{\Delta_d}{s} \quad , \quad \tau_o \equiv \frac{C_o}{g_o} \quad (4.17)$$

Assuming $\Delta_d \cdot k_s^N \gg E_{OH}$ results in $T_c \ll \tau_o$, and hence (4.17) can be simplified into,

$$Y(s) \approx \Delta_d / (s^{N+1} \tau_u^N) \Rightarrow y(t) \approx \frac{\Delta_d}{\tau_u^N} \cdot \frac{1}{N!} \cdot t^N \quad (4.18)$$

and the resolution – speed trade-off becomes,

$$\Delta_d \cdot \left(\frac{T_c}{\tau_u} \right)^N \approx N! \cdot E_{OH} \quad (4.19)$$

The curves corresponding to $N = 2$ and $N = 5$ in Fig. 4.11 illustrate this trade-off for $E_{OH} = 1V$. Note that the multi-step architecture obtains smaller values of T_c than the single-step architecture. For instance, for $\tau_u = 10ns$, $E_{OH} = 1V$ and $\Delta_d = 10mV$ (4.19) yields $T_c \approx 141ns$ for $N = 2$, $T_c \approx 65ns$ for $N = 5$, and $T_c \approx 67ns$ for $N = 8$ – always smaller than for the single-step case.

Fig. 4.12 is a zoom of the resolution – speed curves for the multi-step comparator. It shows that, for each Δ_d , there is an optimum value of N that minimizes T_c . For $\Delta_d > (10^{-3} \cdot E_{OH})$ this optimum number is given by [9],

$$N_{\text{opt}} \approx 1.1 \cdot \ln \left(\frac{E_{OH}}{\Delta_d} \right) + 0.79 \quad (4.20)$$

For instance, for $\Delta_d \approx 10^{-2} \cdot E_{OH}$ maximum speed is achieved by using $N = 6$. Using either less or more stages in the cascade yields slower operation.

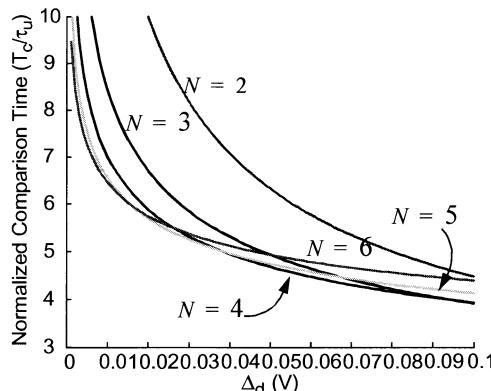


Figure 4.12: Zoom of resolution-speed trade-off for multi-step comparators.

4.3.3 Regenerative Comparators

Let us first consider Fig. 4.6(a); assume the same conditions as in the previous section, and use the model of Fig. 4.6(b) around the quiescent point. Assuming further that g_{mpf} is larger than $g_o = 1/r_o$ and defining $g_{oeq} \equiv g_{mpf} - g_o$ yields,

$$y(t) = \Delta_d \cdot \frac{g_m}{g_{oeq}} \cdot \left(e^{\frac{t\frac{g_{oeq}}{C_o}}{\tau_u}} - 1 \right) ; \quad k_d(t) = \frac{g_m}{g_{oeq}} \cdot \left(e^{\frac{t\frac{g_{oeq}}{C_o}}{\tau_u}} - 1 \right) \quad (4.21)$$

By comparing these expressions with (4.11) and (4.14) and/or (4.15), we note that g_{oeq} contains a term which is proportional to a transconductance ^{†13}. We can assume that $g_{oeq} \approx g_{mpf}$ and consider that $g_{mpf} = g_m$ so that,

$$\Delta_d \cdot e^{\frac{T_c}{\tau_u}} \approx E_{OH} \quad (4.22)$$

Due to the exponential dependence, this equation anticipates much smaller values of T_c than those obtained from Eqs. (4.14) and (19). This is confirmed by the curve labelled as $R = 1$ in Fig. 4.11; the meaning of the parameter R will be explained below. In any case, the set of curves in Fig. 4.11 confirm that regenerative comparators feature faster operation speed than either single-step or multi-step comparators.

Note that, except for the influence of second-order effects, the operation described above is valid no matter how small the input signal magnitude is; only the input sign is meaningful. It means that ideally regenerative comparators might be capable to build infinitely large dynamic gain values – a feature neither shared by single-step nor by multi-step comparators whose maximum dynamic gain is smaller than the static one. In either case, as Δ_d decrease, T_c increases according to,

$$\frac{T_c}{\tau_u} \approx \ln\left(\frac{E_{OH}}{\Delta_d}\right) \quad (4.23)$$

which for $\Delta_d \rightarrow 0$ yields $T_c \rightarrow \infty$.

The previous analysis can be extended to a latch by using the small-signal model of Fig. 4.9(a), which assumes full symmetry (equal positive and negative parameters) and can be represented by the following state equation,

13. For any OTA, the transconductance g_m is usually much larger than the output conductance g_o .

$$\begin{aligned} C_o \cdot \frac{dy_+}{dt} &= -g_o \cdot y_+ - g_m \cdot y_- + g_{m_{in}} \cdot x_+ \\ C_o \cdot \frac{dy_-}{dt} &= -g_m \cdot y_+ - g_o \cdot y_- + g_{m_{in}} \cdot x_- \end{aligned} \quad (4.24)$$

By combining the two equations above, we have

$$C_o \cdot \frac{d(y_+ - y_-)}{dt} = g_m \cdot (y_+ - y_-) - g_o \cdot (y_+ - y_-) + g_{m_{in}} \cdot (x_+ - x_-) \quad (4.25)$$

In the right-hand side of this expression, the first term accounts for the positive feedback effect, the second one for the negative feedback, and the last term represents the input contribution. Further, consider $(g_m/g_o) \gg 1$. Then, assuming that the circuit is initialized at $t = 0$, so that $y(0) = y_+(0) - y_-(0) = 0$, and that a differential input step of value $\Delta_d \equiv x_+ - x_-$ is applied at this time instant, the differential output waveform can be approximated by,

$$y(t) \equiv y_+(t) - y_-(t) \approx \Delta_d \cdot \frac{g_{m_{in}}}{g_m} \cdot e^{\frac{t}{C_o}} \equiv \Delta_d \cdot \frac{g_{m_{in}}}{g_m} \cdot e^{\frac{t}{\tau_u}} \quad (4.26)$$

A similar equation is found when the latch is driven during the reset phase, establishing a voltage unbalance $y(0) = y_+(0) - y_-(0) \equiv \Delta_d$, and no input is applied during the comparison phase. In such a case,

$$y(t) \approx \Delta_d \cdot e^{\frac{t}{\tau_u}} \quad (4.27)$$

and the associated resolution – speed trade-off is represented by (4.22). On the other hand, from (4.26) the following resolution – speed trade-off is found,

$$\Delta_d \cdot e^{\frac{T_c}{\tau_u}} \approx E_{OH} \cdot \frac{g_m}{g_{m_{in}}} \quad (4.28)$$

Fig. 4.11 illustrates this trade-off for different values of $R \equiv g_m/g_{m_{in}}$.

The analysis above can be easily extended to pre-amplified regenerative comparators. Actually, (4.28) is representative of the behavior of the circuits belonging to this class. Bear in mind that this equation has been obtained for Fig. 4.9(a), which is a first-order model of the pre-amplified regenerative comparator of Fig. 4.10(b). By comparing this equation with (4.22), it can be concluded that pre-amplification relaxes the trade-off proportionally to the $g_{m_{in}}/g_m$ ratio, as anticip-

pated. This result can be extended to the circuit of Fig. 4.10(a), where the improvement will be proportional to A_{in} .

However, this is an oversimplified approach because it ignores the dynamic behavior of the circuit used for pre-amplification. More realistic calculations can be made by considering that the latch starts operating with some delay with respect to the pre-amplifier. Hence, by assuming that this delay is fixed, say of value T_{CP} , and the pre-amplifier behaves as a single-step comparator, the following resolution – speed trade-off is derived,

$$\Delta_d \cdot \frac{T_{CP}}{\tau_u} \cdot e^{\frac{T_c}{\tau_u}} \approx E_{OH} \quad (4.29)$$

4.4 ON THE IMPACT OF THE OFFSET

Previous calculations have assumed that comparators are perfectly balanced at the quiescent point so that even arbitrarily small input signals drive the comparator output in the correct direction. However, in practice, there are different errors degrading such an ideal picture. Among these errors, the lack of symmetry of the comparator circuitry results in an equivalent input offset voltage [6]. This offset sets a limitation on the minimum achievable comparator resolution; input signals whose amplitude is smaller than the input offset voltage will not be properly detected by the comparator.

The offset has two different components. *Deterministic offset* is due to asymmetries of the comparator circuit structure itself; for instance, the FDOTAC structure of Fig. 4.2(c) is symmetric, while the AOTAC structure formed of Fig. 4.2(a) is asymmetric. Consequently, the output voltage at the quiescent point Y_Q will be typically different from zero, thus making $E_{os} = Y_Q/k_s$. However, because Y_Q could be, at worst, of the same order of magnitude as E_{OH} , the role played by the deterministic offset component is similar to that of the static gain. On the other hand, *random offset* contemplates asymmetries caused by random deviations of the transistor sizes and technological parameters, and it is observed in both asymmetrical and symmetrical circuit topologies. These deviations make nominally identical transistors become mismatched, the amount of mismatch being inversely proportional to the device area and directly proportional to the distance among them. Assuming that the area-dependent contribution dominates mismatch among devices^{†14}, the statistical variances of the deviations on the zero-bias threshold voltage and the intrinsic large-signal transconductance density between two nominally identical MOSTs are formulated as [22],

$$\sigma^2(\Delta V_{T0}) \approx \frac{\alpha_{V_{T0}}^2}{W \cdot L} \quad , \quad \sigma^2\left(\frac{\Delta \beta_0}{\beta_0}\right) \approx \frac{\alpha_{\beta_0}^2}{W \cdot L} \quad (4.30)$$

where W and L are the channel width and length of the MOSTs, $\alpha_{V_{T0}}^2$ and $\alpha_{\beta_0}^2$ are technological constants^{†15}, and it has been assumed that parameter variations of individual transistors are non-correlated.

With (4.30) and the large-signal MOST model of Appendix 6, a simple variational analysis can be performed to estimate the random offset voltages of single-step topologies in terms of transistor dimensions, technological parameters and biasing conditions. It is worth noting that results are also extensible to multi-steps comparators because, as indicated in (4.6), the overall offset voltage is dominated by that of the front-end stage.

For the sake of conciseness, only the AOTAC structure, Fig. 4.2(a), will be considered herein. In this topology, contributions to the random offset voltage arise from mismatches in the differential input pair, formed by transistors M_{ND} , and imbalance in the active load circuitry, formed by transistors M_{PL} . Such contributions can be estimated by calculating the input voltage required to compensate any current unbalance through the comparator branches as a result of transistor mismatch. Considering the differential pair mismatch alone, the variance of its contribution to the random offset is found to be,

$$\sigma^2(E_{os})|_{DP} = \frac{2 \cdot \alpha_{V_{T0DP}}^2}{W_{DP} \cdot L_{DP}} + \frac{I_B \cdot n_{DP}}{4 \cdot \beta_{0DP}} \cdot \frac{\alpha_{\beta_{0DP}}^2}{W_{DP}^2} \quad (4.31)$$

and, in case of the active load alone, we have,

$$\begin{aligned} \sigma^2(E_{os})|_{AL} &= \beta_{0AL} \frac{2 \cdot \beta_{0AL} \cdot n_{DP}}{n_{AL}} \cdot \frac{L_{DP}}{W_{DP}} \cdot \frac{\alpha_{V_{T0AL}}^2}{L_{AL}^2} + \\ &+ \frac{I_B \cdot n_{DP}}{4 \cdot \beta_{0DP}} \cdot \frac{L_{DP}}{W_{DP}} \cdot \frac{\alpha_{\beta_{0AL}}^2}{W_{AL} \cdot L_{AL}} \end{aligned} \quad (4.32)$$

Assuming that both contributions are uncorrelated, the variance of the random offset voltage of the AOTAC structure is, thus, obtained as

-
- 14. The distance contribution to mismatch accounts for the effect of process-parameter gradients on the wafer and can be largely attenuated through proper layout techniques (e.g., common-centroid structures).
 - 15. Typical values in a $0.5\mu\text{m}$ technology are $\alpha_{V_{T0}}^2 \approx 10^{-5}\text{V}^2\mu\text{m}^2$ and $\alpha_{\beta_0}^2 \approx 10^{-4}\mu\text{m}^2$.

$$\begin{aligned}\sigma^2(E_{os}) = & \left. \sigma^2(E_{os}) \right|_{DP} + \left. \sigma^2(E_{os}) \right|_{AL} = \frac{k_1}{W_{DP} \cdot L_{DP}} + \\ & + \frac{k_2 \cdot I_B}{W_{DP}^2} + \frac{k_3 \cdot L_{DP}}{W_{DP} \cdot L_{AL}^2} + \frac{k_4 \cdot L_{DP}}{W_{DP}} \cdot \frac{I_B}{W_{AL} \cdot L_{AL}}\end{aligned}\quad (4.33)$$

where parameters

$$\begin{aligned}k_1 &= 2 \cdot \alpha_{V_{TODP}}^2 & k_2 &= \frac{n_{DP}}{4 \cdot \beta_{0DP}} \alpha_{\beta_{0DP}}^2 & k_3 &= \frac{2 \cdot \beta_{0AL}}{\beta_{0DP}} \cdot \frac{n_{DP}}{n_{AL}} \cdot \alpha_{V_{TOAL}}^2 \\ k_4 &= \frac{I_B \cdot n_{DP}}{4 \cdot \beta_{0DP}} \cdot \alpha_{\beta_{0AL}}^2\end{aligned}\quad (4.34)$$

are constant for a given technology.

This result can be exploited to examine the resolution – speed trade-off in terms of design variables and technological parameter variations as follows. Let us define the comparator *dynamic resolution* as $\xi_d = |E_{os}| + \Delta_d$, in consonance with the static resolution defined in (4.2). Then, each of the contributing terms to ξ_d is expressed in terms of the physical parameters involved in the topology. On the one hand, it can be conveniently assumed that $|E_{os}| \approx 3\sigma(E_{os})$, where $\sigma(E_{os})$ is obtained by square-rooting (4.33). On the other, taking into account (4.14) and Table 4.1.; we can express Δ_d as

$$\Delta_d \approx \frac{E_{OH} \cdot \tau_u}{T_c} = \frac{E_{OH}}{T_c} \cdot \frac{[C_L + v_{DP} \cdot W_{DP} \cdot C_{GD0DP} + v_{AL} \cdot W_{AL} \cdot C_{GD0AL}]}{\sqrt{2 \cdot \frac{\beta_{0DP}}{n_{DP}} \cdot \left(\frac{W}{L}\right)_{DP} \cdot I_B}} \quad (4.35)$$

Hence, ξ_d can be defined as a function of T_c , and also of the device dimensions and biasing. Fig. 4.13 illustrates such a relationship by individually representing ξ_d against the variables implicated. It is worth pointing out that such representations exhibit minima that indicate the values of corresponding variables for which an optimum resolution – speed performance can be obtained.

Let us now briefly consider the effect of asymmetries on the performance of discrete-time regenerative comparators. Spurious differential signals, coupling between latch branches and mismatches between their parameters preclude correct amplification of small Δ_d values. Their influence can be assessed by studying the equilibrium points, eigenvalues and eigenvectors of the state equation,

$$\begin{aligned}(C_{o+} + C_c) \cdot \frac{dy_+}{dt} &= -g_{o+} \cdot y_+ - g_{m+} \cdot y_- + g_{m_{in+}} \cdot x_+ + g_{m+} \cdot \frac{E_{os}}{2} + C_c \cdot \frac{dy_-}{dt} \\ (C_{o-} + C_c) \cdot \frac{dy_-}{dt} &= -g_{m-} \cdot y_+ - g_{o-} \cdot y_- + g_{m_{in-}} \cdot x_- - g_{m-} \cdot \frac{E_{os}}{2} + C_c \cdot \frac{dy_+}{dt}\end{aligned}\quad (4.36)$$

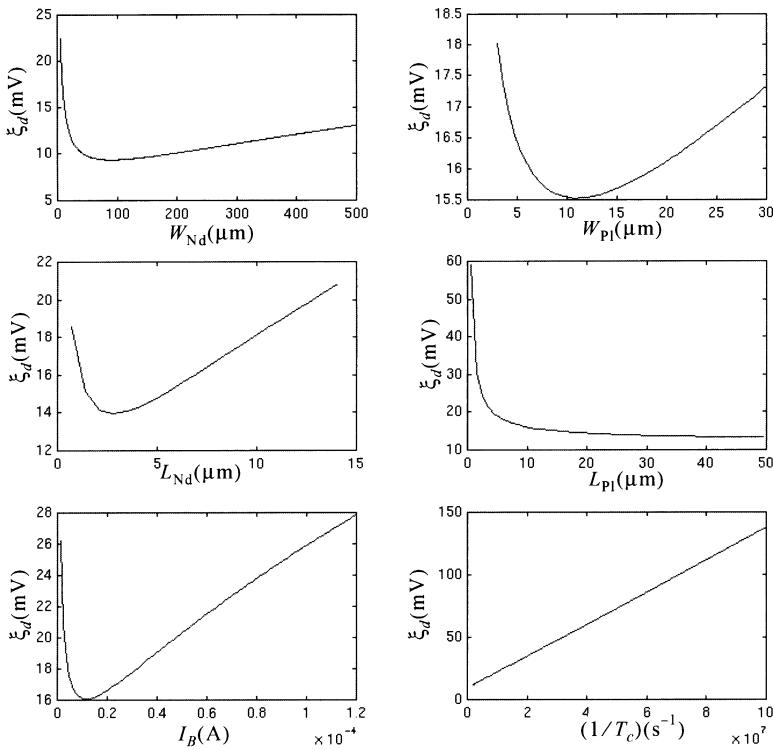


Figure 4.13: Illustration of the resolution - speed trade-off in terms of device dimensions and biasing.

where C_c must be added in Fig. 4.9(a) in order to model the coupling between the two output nodes. This is out of the scope of this chapter and only some remarks will be given.

Note from (4.36) that the influence of the offset voltage E_{os} is similar to that observed in single-step and multi-step comparators. However, asymmetries between transconductances g_{m+} and g_{m-} , as well as between capacitors C_{o+} and C_{o-} , produce much larger errors for regenerative comparators than for single-step and multi-step comparators. It can be shown that the error depends on the common-mode value of the input signal $x_{cm} = (x_+ + x_-)/2$ [12]. For zero common-mode, (4.36) does not reveal any limitation on Δ_d . However, as the common-mode increases up to half of the swing range, $|\Delta_d|$ has to be larger than $\approx 30\text{mV}$ for correct codification of the input signal polarity. This value increases up to $\approx 50\text{mV}$ if 10% mismatches are considered for transconductances and capacitances. Clearly, this poses a hard constraint on the comparator resolution –

not shared by either single-step or multi-step comparators. As it was already mentioned, this problem is overcome by placing a pre-amplifier in front of the regenerative core and using offset-compensation techniques.

4.5 OFFSET-COMPENSATED COMPARATORS

Although input offset voltage can be compensated through transistor sizing, these techniques can hardly obtain offsets lower than a few millivolts – not small enough for many practical applications. This drawback can be overcome by adding offset-cancellation circuitry; thus residual offset values as smaller as 0.1 mV can be obtained [23]. Out of the different offset-cancellation strategies – component trimming, error integration and feedback through an offset nulling port, ... –, here we focus only on the use of dynamic biasing.

4.5.1 Offset-Compensation Through Dynamic Biasing

A simple, yet efficient offset correction technique uses dynamic self-biasing in order to, first, extract and store the offset, and then annul its influence [24] [25]. Figs. 4.14(a) and (c) show the circuit implementation of this technique for single-ended and fully-differential OTAs, respectively. As it can be seen, both circuits operate in discrete-time, under the control of a clock with two non-overlapped phases as indicated in Fig. 4.14(a). When the clock phase φ_a is in the low state, and correspondingly φ_r is in the high-state, switches controlled by φ_r are ON, the others are OFF, and the comparator is *self-biased – reset phase*. In this phase, referring to Fig. 4.14(a), the output voltage evolves towards a steady-state value,

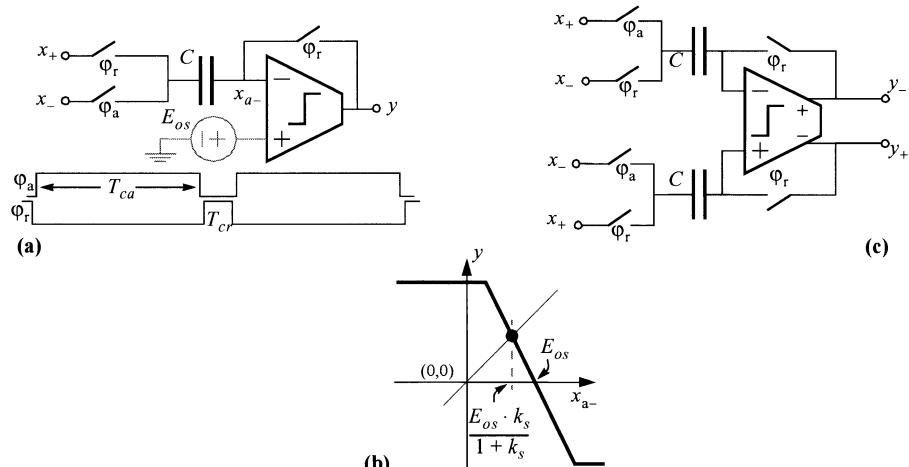


Figure 4.14: Offset cancellation in single-step voltage comparators.

$$x_{a-}|_r = E_{os} \cdot (1 + k_s^{-1}) \quad (4.37)$$

defined by the intersection of the amplifier transfer characteristic and the bisecting line $y = x_{a-}$, as Fig. 4.14(b) illustrates. Provided that the *reset* interval is long enough for the transient to vanish, this value is stored at capacitor C , so that,

$$v_{C_r} = x_+ - x_{a-}|_r \quad (4.38)$$

Note that for $k_s \gg 1$ it yields,

$$x_{a-}|_r \approx E_{os} \quad (4.39)$$

Hence, during the reset phase the plate of the capacitor connected to the non-inverting terminal of the OTA samples a voltage very close to the offset.

On the other hand, when φ_a is in the high state, and correspondingly φ_r is in the low-state, switches controlled by φ_r are OFF, the others are ON, and C keeps its charge – comparison, or active phase. In this phase, the comparator input $x_a \equiv x_{a+} - x_{a-}$ evolves to a steady-state,

$$x_a = E_{os} - (x_- - v_{C_r}) = E_{os} - x_{a-}|_r + (x_+ - x_-) \quad (4.40)$$

where the offset is subtracted from its previously sampled voltage. This results in the following static resolution expression,

$$\xi_S \approx \frac{|E_{os}|}{1 + k_s} + \frac{E_{OH}}{k_s} = \frac{|E_{os}|}{1 + g_m/g_o} + E_{OH} \cdot \frac{g_o}{g_m} \quad (4.41)$$

which shows that the offset error is attenuated by a factor $1 + k_s$ as compared to the uncompensated comparator – see (4.2).

The circuits in Figs. 4.14 correct the offset by sampling-and-holding its value at the comparator input node. Alternatively, the offset can be compensated through a sample-and-hold operation at the output node. In the circuit of Fig. 4.15(a) such an operation is realized in the voltage domain. During the reset phase, the output offset voltage $k_s \cdot E_{os}$ is stored at the node y_a , while the output node is tied to the analog ground. Then, during the active phase, the inputs are applied yielding $y_a = k_s \cdot E_{os} + k_s \cdot (x_+ - x_-)$ and since no current is circulating through the capacitor C , $y = k_s \cdot (x_+ - x_-)$. Obviously, for proper offset correction k_s must be low enough to guarantee that the OTA remains operating within its linear region, i.e., that its output is not saturated during the reset phase. The circuit of Fig. 4.15(b) employs a different offset storage mechanism that overcomes this

problem. There, the output offset current, instead of the voltage, is stored during the reset phase, and then subtracted during the comparison phase. Current storage is realized by transistor M_{OS} , which operates as a current memory. Note that during the reset phase this transistor operates as a nonlinear resistor, setting the OTA output node at low impedance. This yields a significant attenuation of the voltage gain during the reset phase, thereby reducing the excursions of the output voltage and guaranteeing operation within the OTA linear region.

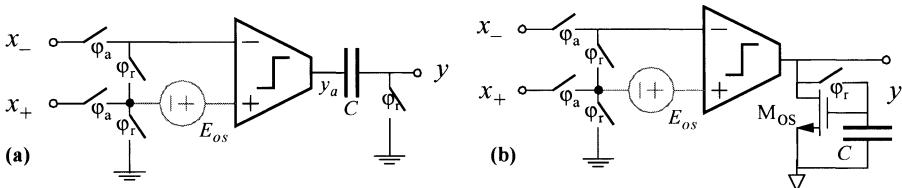


Figure 4.15: Offset compensation through storage at the output node: (a) Voltage storage; (b) Current storage.

4.5.2 Offset Compensation in Multi-Step Comparators

Dynamic self-biasing can be also applied to cancel out the offset of multi-step comparators. However, unless compensation circuitry is added, the high-order dynamics will cause instabilities when a direct feedback connection is established between the overall output node and the negative input – similar to the problem found in two-stage opamps [6] [26]. Instabilities can be avoided by making each stage to store its own offset, as shown in Fig. 4.16, so that only second-order offset terms generated at the different stages remain. These offset terms can be further attenuated through a proper sequential timing of the switches used for self-biasing. The inset of Fig. 4.16 shows this timing. Note that the stages are switched ON at different, consecutive time instants. Consequently, the residual offset of each stage is stored at the input capacitor of the next one while this latter remains grounded, and hence the output remains unaltered. In this way only the residual offset - see next section - of the last stage $|E_{osdN}|$ contributes to the output. Since this offset is amplified only by the last stage itself, while the signal is amplified by all the stages, the following expression is obtained for the static resolution,

$$\xi_s \approx \frac{|E_{osdN}|}{k_s^{N-1}} + \frac{E_{OH}}{k_s^N} \quad (4.42)$$

Offset compensation applies also to multi-step topologies formed by cascading a pre-amplifier and a latch. It is illustrated through the circuit depicted in

Fig. 4.16(b) consisting of the cascade of a self-biased single-step comparator and a self-biased latch. [9] [27].

4.5.3 Residual Offset and Gain Degradation in Self-Biased Comparators

There are several second-order phenomena that modify the voltage stored at node x_{a-} in Fig. 4.14(a), and consequently degrade the static resolution of self-biased comparators. The two most important effects take place during the ON \rightarrow OFF transition of the reset feedback switch, namely: a) the feedthrough of the clock signal that controls this switch, and b) the injection of its channel charge. They make the voltage stored at node x_{a-} experience a finite jump during the ON \rightarrow OFF transition so that its value at the active phase differs from that stored during the reset phase, i.e.

$$x_{a-}|_a \approx x_{a-}|_r - \Delta x_{a-} \quad (4.43)$$

Also, during the active phase this value continues degrading due to the switch leakage current, I_{leak} . Fig. 4.17 shows a simplified model to evaluate all these degradations. In addition to the nominal capacitor C , this model includes a parasitic capacitor, C_{a-} , between node x_{a-} and ground, and a parasitic capacitor, C_{ov} , between node x_{a-} and the control terminal of the feedback switch. Analysis using this model provides the following expression for the static resolution,

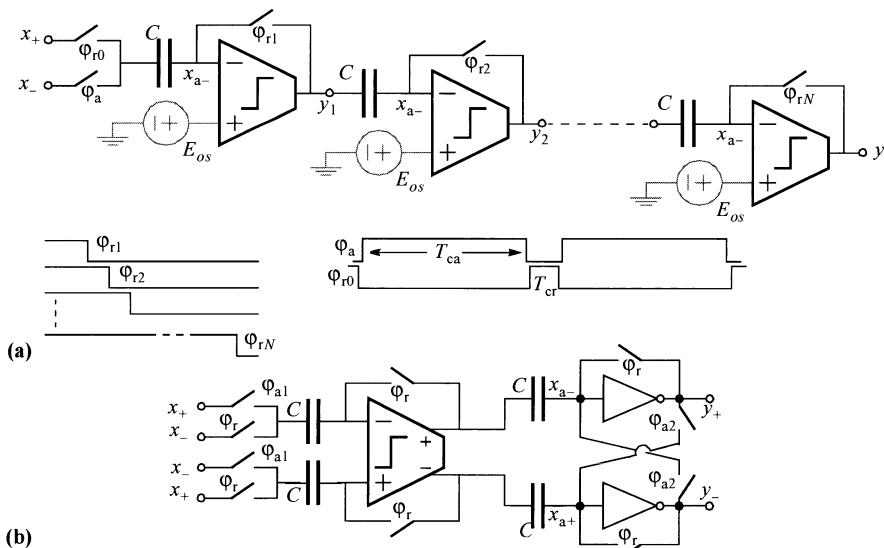
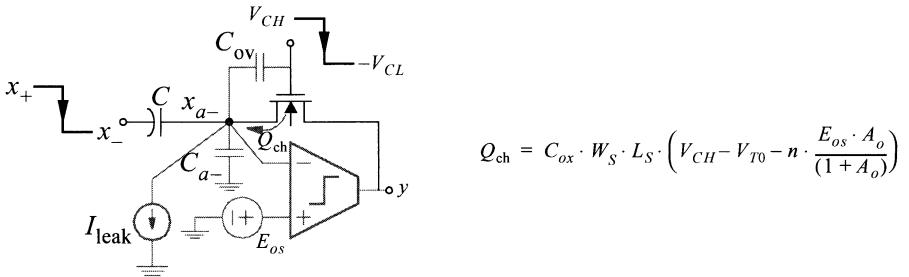


Figure 4.16: Offset Compensation in multi-step voltage comparators.



$$Q_{ch} = C_{ox} \cdot W_S \cdot L_S \cdot \left(V_{CH} - V_{T0} - n \cdot \frac{E_{os} \cdot A_o}{(1 + A_o)} \right)$$

Figure 4.17: Showing the feedthrough and the leakage current influence.

$$\xi_s \approx |V_{CH} + V_{CL}| \cdot \frac{C_{ov}}{C} + \frac{|Q_{ch}|}{C} + \frac{|I_{leak}|}{C} \cdot t + \frac{|E_{os}|}{\alpha_C \cdot (1 + k_S)} + \frac{E_{OH}}{\alpha_C \cdot k_S} \quad (4.44)$$

where $\alpha_C = C / (C + C_{ov} + C_{a-})$; V_{CH} and V_{CL} are, respectively, the high and low state levels of the clock signal; Q_{ch} is the charge accumulated in the switch channel when it is ON (during the reset phase), and t is the time measured from the instant when the ON \rightarrow OFF transition happens. Note that (4.44) shows a residual offset term,

$$|E_{osd}| = |V_{CH} + V_{CL}| \frac{C_{ov}}{C} + \frac{|Q_{ch}|}{C} + \frac{|I_{leak}|}{C} t \quad (4.45)$$

that is not attenuated by the comparator gain. If capacitance C is chosen very small, $|E_{osd}|$ may become larger than the original offset $|E_{os}|$. Also, small values of this capacitance may result in small values of α_C , thus making the last term in (4.44) increase, and producing additional resolution degradation.

The above equation shows that resolution degradations caused by residual offset are attenuated by increasing C and decreasing C_{ov} , Q_{ch} and I_{leak} . These three latter parameters are related to the ON resistance of the transistor switch. On the one hand,

$$Q_{ch} = \frac{1}{\mu} \cdot \frac{L_S^2}{R_{ON}} \quad (4.46)$$

On the other, both C_{ov} and I_{leak} are proportional to the width W_S of the transistor switch, and hence inversely proportional to R_{ON} . Thus, the measures for reducing the residual offset make this resistance increase; consequently, the time constants increase as well and a resolution – speed trade-off appears.

4.5.4 Transient Behavior and Dynamic Resolution in Self-Biased Comparators

During the active phase of self-biased comparators, the transient response follows an exponential behavior similar to (4.11); the difference is that the static resolution, k_s , is now attenuated by α_C – see (4.44) –. Hence, the resolution – speed trade-off discussed in relation to (4.15) also applies in this case.

On the other hand, another trade-off stems from transient during the reset phase, related to the onset of an additional residual offset component. The dynamic behavior within the reset phase can be calculated using the model of Fig. 4.18(a). Two different transients are observed. First of all, there is a very fast charge redistribution transient, dominated by the ON resistances of the switches. The output value $y(0)$ at the end of this transient will be, at worst, equal to one of the saturation levels. Let us assume $y(0) = E_{OH}$. From this value, the output evolves towards the steady-state located at

$$E_{os} \cdot (1 + k_s^{-1})^{-1} \quad (4.47)$$

through a second transient which is dominated by the comparator dynamics. Fig. 4.18(b) provides a global view of this second transient. It consists of a linear part, where the transconductor is in the saturation region and y evolves from $y(0)$ to δ_m with a fixed slew-rate δ_m / τ_u , followed by an exponential part with time constant

$$\tau_{ur} = \frac{(C + C_{a-} + C_o)}{(g_m + g_o)} \approx \frac{C}{g_m} \equiv \tau_u \quad (4.48)$$

Thus, the *reset time*, T_r , needed to reach a final Δx_{a-} above the steady-state

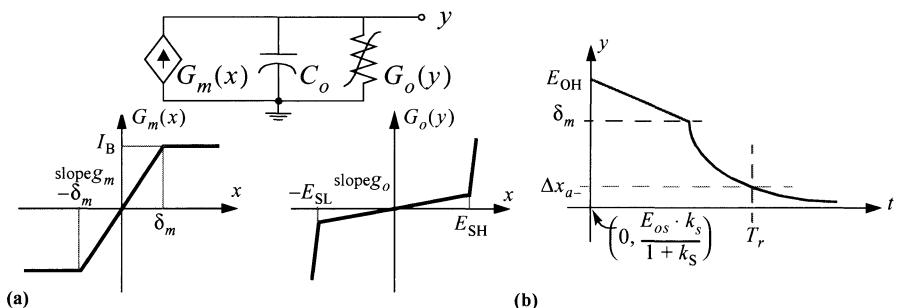


Figure 4.18: Non-linear model for the one step comparator reset phase.

value is given by,

$$T_r \approx \frac{E_{OH} - \delta_m}{\delta_m} \cdot \tau_u + \tau_u \cdot \ln\left(\frac{\delta_m}{\Delta x_{a-}}\right) \quad (4.49)$$

Note that Δx_{a-} will remain as a residual offset after cancellation.

This equation shows another resolution – speed trade-off. The smaller Δx_{a-} , and hence the error, the larger the time T_r needed for reset. Considering the typical values of $k_s = 2 \times 10^3$, $\tau_u = 10\text{ns}$, $E_{OH} = 1\text{V}$ and $\delta_m = 250\text{mV}$, (4.49) obtains $T_r \approx 8.5\mu\text{s}$ for 1mV residual offset. Note that this time is shorter than the amplification time ($T_c \approx 14\mu\text{s}$) required to obtain $\Delta_d = 1\text{mV}$ from (4.15).

At this point new error sources should be considered for completeness. Particularly, noise is an issue in the case of self-biased comparators, although it is usually less important than offset for those comparators which are not self-biased. For considerations related to noise readers are referred to other chapters of this book.

4.6 APPENDIX I: SIMPLIFIED MOST MODEL

MOS transistors exhibit different operation depending on the current and voltage levels. Through the article we consider the MOST model only under *strong* channel inversion, and describe its first-order behavior using a model with four parameters, namely: zero-bias *threshold voltage* V_{T0} , *slope factor* n , *intrinsic transconductance density* β_0 , and *equivalent Early voltage* V_A [28]. Two subregions are considered within strong inversion:

- *Triode (or ohmic) region*. In this regime, the source and drain voltages V_S , V_D remains below $V_p = (V_G - V_{T0})/n$, where V_G is the gate voltage (all voltages are referred to the local substrate). The drain current takes the form,

$$I_D = 2\beta_0 \cdot \frac{W}{L} \cdot \left[V_G - V_{T0} - \frac{n}{2} \cdot (V_D + V_S) \right] \cdot (V_D - V_S) \quad (4.50)$$

where W/L is the aspect ratio of the transistor.

- *Saturation region*. Assuming forward operation, this regime is reached when $V_S < V_p < V_D$ and the drain current amounts,

$$I_D = \beta_N \cdot (V_G - V_{T0} - nV_S)^2 \cdot \left[1 + \frac{V_D - V_p}{V_A} \right] \quad (4.51)$$

where $\beta \equiv \frac{\beta_0}{n} \cdot \frac{W}{L}$.

4.7 REFERENCES

- [1] B. Razavi, *Principles of Data Conversion System Design*. IEEE Press, 1995.
- [2] F. Medeiro, B. Pérez-Verdú and A. Rodríguez-Vázquez, *Top-Down Design of High-Performance Sigma-Delta Modulators*. Kluwer, 1999.
- [3] A. Rodríguez-Vázquez, M. Delgado-Restituto and F. Vidal, “Synthesis and Design of Nonlinear Circuits”. Section VI.32 in *The Circuits and Filters Handbook* (edited by W.K. Chen). CRC Press, 1995.
- [4] L.E. Larson (editor), *RF and Microwave Circuit Design for Wireless Communications*. Artech House, 1996.
- [5] A. Cichocki and R. Unbehauen, *Neural Networks for Optimization and Signal Processing*. John Wiley, 1993.
- [6] K.R.Laker and W.M.C. Sansen, *Design of Analog Integrated Circuits and Systems*, McGraw-Hill, 1994.
- [7] J. F. Duque-Carrillo, “Control of the Common-Mode Component in CMOS Continuous-Time Fully Differential Signal Processing”. *Analog Circuits and Signal Processing*, Vol. 4, pp. 131-140, 1993.
- [8] K. Bult and G.J.G.M. Geelen, “A Fast Settling CMOS Opamp for SC Circuits with 90-dB DC Gain”. *IEEE Journal of Solid-State Circuits*, Vol. 25, pp. 1379-1384, 1990.
- [9] B. Razavi and B.A Wooley, “Design Techniques for High-Speed, High-Resolution Comparators”. *IEEE Journal of Solid-State Circuits*, Vol. 27, pp. 1916-1926, 1992.
- [10] J.T. Wu and B.A. Wooley, “A 100-MHz Pipelined CMOS Comparator”. *IEEE J. Solid-State Circuits*, Vol. 23, pp. 1379-1385, 1988.
- [11] S. Dhar and M.A. Franklin, “Optimum Buffer Circuits for Driving Long Uniform Lines”. *IEEE J. Solid-State Circuits*, Vol. 26, No. 1, pp. 32-40, 1991.
- [12] A. Rodríguez-Vázquez, R. Domínguez-Castro, F. Medeiro and M. Delgado-Restituto, “High Resolution CMOS Current Comparators: Design and Applications to Current-Mode Function Generation”. *Analog Integrated Circuits and Signal Processing*, Vol. 7, No. 2, pp. 149-165, March 1995.
- [13] L.O. Chua, C.A. Desoer and E.S. Kuh, “Linear and Nonlinear Circuits”. McGraw-Hill, New-York 1987.
- [14] A. Rodríguez-Vázquez, M. Delgado-Restituto, and R. Domínguez-Castro: “Comparator Circuits”.at *Wiley Encyclopedia of Electrical and Electronics Engineering* (edited by John G. Webster), ISBN 0-471-13946-3, Vol. 3, pp. 577-600, John Wiley & Sons, New-York 1999.
- [15] L.O. Chua, C. A. Desoer and E. S. Kuh, *Linear and Nonlinear Circuits*. McGraw-Hill 1987.
- [16] L.A. Glasser and D.W. Dobberpuhl, *The Design and Analysis of VLSI Circuits*. Addison Wesley, 1985.
- [17] B. Ginetti, P.G.A. Jespers and A. Vandemeulebroecke, “A CMOS 13-b Cyclic RSD A/D Converter”. *IEEE J. Solid State Circuits*, Vol. 27, No. 7, pp. 957-965, 1992.

- [18] B.S. Song, S.H. Lee and M.F. Thompsett, "A 10-b 15-MHz CMOS Recycling Two-Step A/D Converter". *IEEE J. Solid State Circuits*, Vol. 25, No. 6, pp. 1328-1338, 1990.
- [19] G.M. Yin, F. Opt't Eyende, and W. Sansen, "A High-Speed CMOS Comparator with 8-b Resolution", *IEEE J. Solid-State Circuits*, Vol. 27, pp.208-211, 1992.
- [20] A. Yukawa, "A Highly Sensitive Strobed Comparator", *IEEE J. Solid-State Circuits*, Vol. SC-16, pp. 109-113, 1981.
- [21] A. Yukawa, "A CMOS 8-Bit High-speed A/D Converter IC", *IEEE J. Solid-State Circuits*, Vol. SC-20, pp. 775-779, 1985.
- [22] M.J.M. Pelgrom, A.C.J. Duinmaijer, and A.D.G. Welbers, "Matching Properties of MOS Transistors". *IEEE J. Solid-State Circuits*, Vol. 24, pp.1433-1440, 1989.
- [23] J.H. Atherton and J.H. Simmonds, "An Offset Reduction Technique for Use with CMOS Integrated Comparators and Amplifiers". *IEEE J. Solid-State Circuits*, Vol. 27, No. 8, pp. 1168-1175, 1992.
- [24] D.J. Allstot, "A Precision Variable-Supply CMOS Comparator". *IEEE J. Solid-State Circuits*, Vol. 17, pp. 1080-1087, 1982.
- [25] Y.S. Yee, L.M. Terman, and L.G. Heller, "A 1 mV MOS Comparator". *IEEE J. Solid-State Circuits*, Vol.13, pp. 63-66, 1978.
- [26] J.K. Roberge, *Operational Amplifiers Theory and Practice*. John Wiley, 1975.
- [27] W.T. Ng and C.A.T. Salama, "High-Speed High-Resolution CMOS Voltage Comparator", *Electronics Letters*, Vol. 22, No. 6, pp.338-339, 1986.
- [28] E.A. Vittoz, "The Design of High-Performance Analog Circuits on Digital CMOS Chips". *IEEE J. Solid-State Circuits*, Vol. 20, pp. 657-665, 1985.

Chapter 5

FOLDING/INTERPOLATING ADCS

Analog Preprocessing Techniques for High-Speed 8-bit ADC

Koen Uyttenhove, J. Vandenbussche, G. Gielen and M. Steyaert

Abstract Designing high-speed flash ADCs in a deep submicron technology requires optimized architectures and building blocks. In this chapter, the design of a high-speed 8 bit converter is presented , following the discussion of analog pre-processing techniques necessary to reduce the power consumption and input capacitance of the converter. Three analog preprocessing techniques will be described, two of which will be used in the design of the 8-bit converter. The systematic design of an 8-bit interpolating/averaging converter will then be described and experimental results will be presented.

5.1 INTRODUCTION

Analog preprocessing techniques are necessary to reduce the power consumption of future high-speed analog to digital converters (ADC). In this chapter two of these techniques are used to implement a high-speed 8-bit, 200-MSample/s ADC. After a general introduction on the different preprocessing techniques used in the design of high-speed analog to digital converters, the systematic design of a high-speed, high-accuracy Nyquist-rate A/D converter will be proposed. The presented design methodology covers the complete flow and is supported by software tools. A generic behavioral model is used to explore the A/D converter's specifications during high-level system design and exploration. The inputs to the coconverter design are the specs of the A/D converter and the technology process. The result is a generated layout and the corresponding extracted behavioral model. The approach has been applied to a real-life test case, where a Nyquist-rate 8-bit 200 MS/s 4-2 interpolating/averaging A/D converter was developed.

5.2 ANALOG PREPROCESSING TECHNIQUES IN ADCS

Flash architectures are typically the simplest and the fastest structures that can be used to implement analog to digital converters. Figure 5.1 presents a block diagram of a N -bit flash converter. The resistive ladder subdivides the converter reference

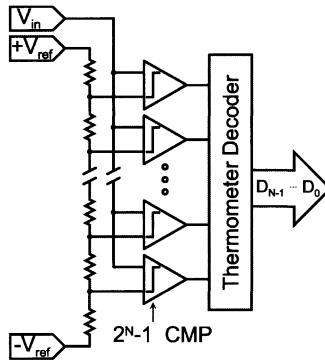


Figure 5.1: Block diagram of the Flash A/D converter.

voltage ($+V_{ref} - V_{ref}$) in a set of 2^N reference voltages, which are compared in parallel with the analog input signal. A logic decoder converts the thermometer code generated by all the comparators into a binary code that approximates the input signal every clock cycle.

Note that the major advantages (simplicity and parallelism) of flash architectures also present its main problem: the number of comparators increases exponentially with the resolution specification, leading typically to a large die area and a high power consumption. Normally, this architecture is only used to implement converters with a resolution less than about 6 bit. Also, the large input capacitance limits the use of this structure for high-speed converters with resolutions above 6 bit.

Therefore, three preprocessing techniques are used in the area of high-speed converters with a resolution of 8 bits and more [Raz 95]:

- Folding techniques
- Interpolating techniques
- Averaging techniques

These three techniques will now be discussed in more detail whereby their advantages and disadvantages will be pointed out.

5.2.1 Folding techniques

Folding is a preprocessing technique, used to reduce the number of comparators in a high-speed ADC. By doing this, power consumption can be decreased [Ven 96]. The folding technique consists in subdividing the input-output characteristic of a typical flash architecture into several parts of equal length which are folded

together (as shown in figure 5.2). To distinguish the exact location of the analog input, the coarse comparison determines in which fold the input is situated and then the fine conversion determines the exact location within this fold. These two operations can be performed in parallel (as shown in figure 5.2).

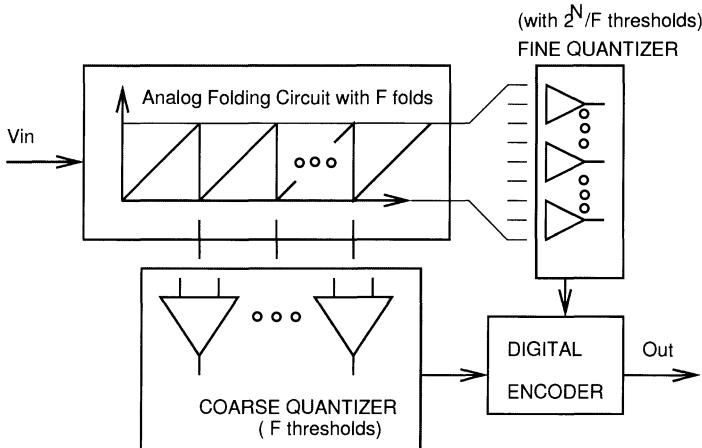


Figure 5.2: Folding Principle.

These folding techniques can still be subdivided into several types of folding: current folding, voltage folding. One of the most important parameters of folding is the number of folds F . This parameter determines the resolution of both the coarse and fine conversion.

Folding A/D converters based on the architecture of figure 5.2 would be possible if simple analog circuits could easily realize the piece-wise linear input-output characteristics indicated. However, because of the discontinuities in the folding transfer function, these types of ideal characteristics are inherently difficult to realize.

A folding A/D converter based upon a periodic, but not piece-wise linear folding function can be developed if the nonlinear function is well behaved which in this context means that the characteristic is periodic and does not saturate. A sinusoidal folding function serves as an example. In this case, the folding characteristic resembles the triangle-wave depicted in figure 5.3; however, the output signal from the folding circuit is not a linear function of the input. Therefore, the fine quantizer must contain thresholds which are not uniformly spaced, but which are located according to an inverse-sine law. This threshold placement would be very difficult to achieve in practice, but a simple alternative lies in generating many sinusoids, uniformly shifted in phase with respect to each other. In this arrangement, the fine quantizer consists of an array of comparators, each with its reference input grounded

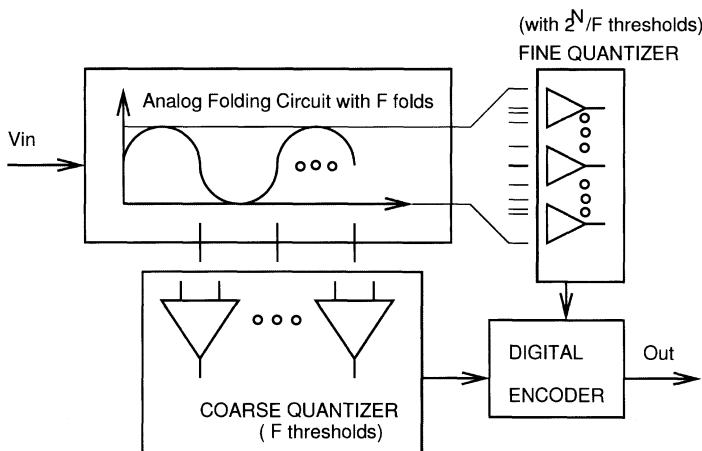


Figure 5.3: Folding Principle with sinewave folding circuit.

so that the quantizer thresholds correspond to the zero-crossings of the sinusoids (see figure 5.4). Since the sinusoids are equally spaced in phase, their zero crossings are equally spaced along the analog input range, and the quantizer thresholds are located correctly. So in fact, instead of level-crossing detection, sinusoidal folding techniques use zero-crossing techniques.

Several CMOS implementations exist for this folding behavior. One of the major problems with folding is the possible limitation of the maximum input frequency. This is because in the folding circuitry the internal frequency is much larger than the input frequency. By using the techniques described above (the zero-crossing detection instead of level detection) this disadvantage is also reduced, thereby making this folding technique a very often used technique to implement medium-resolution analog to digital converters [Fly 96] [Nau 95].

5.2.2 Interpolation Techniques

A second analog preprocessing technique is the interpolation technique. This technique is based upon the flash architecture. In a flash architecture 2^N preamplifiers are put in parallel to compare the input with an equal number of reference voltages. In an interpolating converter some of the preamplifiers are eliminated and replaced by resistive division ("interpolation") between two outputs of other preamplifiers as shown in figure 5.5. By doing so, the number of preamplifiers can be largely reduced and so power can be saved. This technique is only possible because of the non-ideal behaviour of a preamplifier: the input-output characteristic

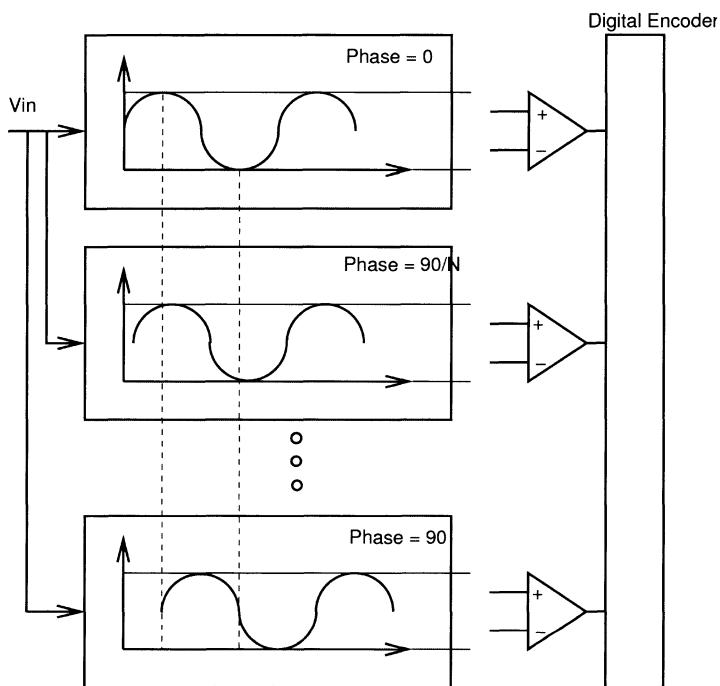


Figure 5.4: Folding Principle with sinewave folding circuit.

of a preamplifier has a linear slope, at least within some input range, allowing the use of interpolation to create additional zero crossings.

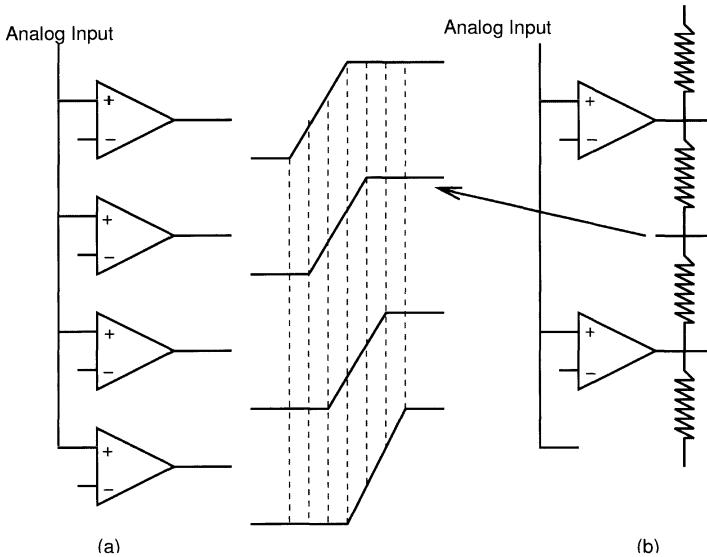


Figure 5.5: Interpolation Principle.

In principle, only two preamplifiers could be used to implement an ADC with this technique. But due to the limited linear range of the preamplifiers (dependent on the gate-overdrive voltage of the preamplifier) only a limited number of preamplifiers can be eliminated.

For fast operation, it is important that the delays to each of the latches are made equal as much as possible. Since the latch comparators have similar input capacitances associated with them, the delays can be made nearly equal by adding extra series resistors, as shown in figure 5.6. These series resistors equalize the impedances seen by each latch comparator looking back into the resistive string [Vdpl].

An additional important benefit of interpolation is the reduction of the differential nonlinearity resulting from the offset of the preamplifiers [Raz 95].

Most of the time this interpolation technique is used together with the already discussed folding technique in a so called folding/interpolating converter. Some folds are created with the folding circuitry but the others are generated by the interpolation technique. With this combination, not only the amount of comparators is reduced but also the number of folding preamplifiers [Raz 95].

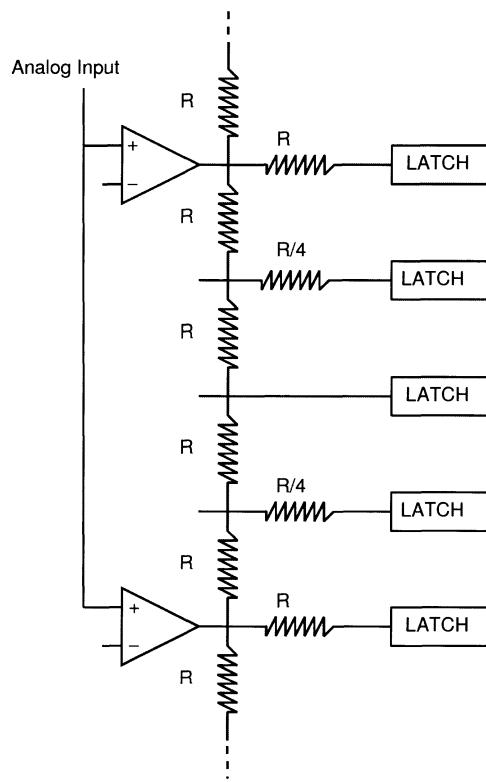


Figure 5.6: Adding series resistances to equalize delay times to latches in interpolating ADCs.

Finally it should be mentioned that circuit techniques other than resistive strings can be used to realize this interpolative approach. In [Roo 93] current mirrors were used to interpolate eight times between comparators. Also capacitive interpolation is a possible implementation technique.

5.2.3 Averaging Techniques

Another technique (similar to the interpolation technique) whereby all the outputs of the preamplifier array are connected with resistors is called the averaging technique. All the fastest low-resolution converters make use of this technique. The averaging technique is illustrated in figure 5.7.

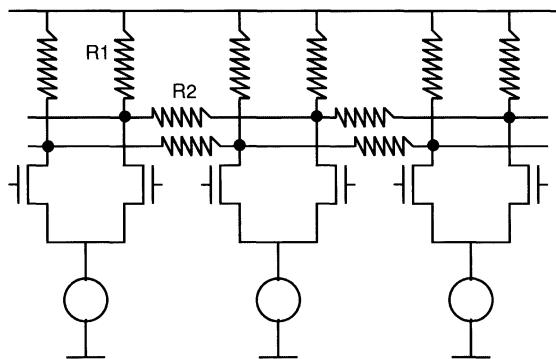


Figure 5.7: Resistive Averaging Implementation.

The original idea of inserting lateral resistors between output nodes of adjacent repetitive cells (e.g. from the comparator preamplifier array) to reduce the effective offset of each cell takes advantage of the difference in impedance seen by the random offset currents and the signal currents [Kat 91]. The improvement is given by the ratio of the signal impedance to the offset impedance. Current sources in place of finite preamp output impedances R_0 's are proposed to compensate for the loss in preamp output impedance so as to maintain the voltage gain. In fact, the original averaging scheme suffers neither loss in voltage gain nor loss in GBW. Even though R_1 's reduce the output impedance seen by each individual preamp, they pass currents from adjacent cells and effectively increase the transconductance of one individual cell, which can compensate for the loss in output impedance. In contrast averaging with infinite R_0 causes problems such as BW degradation, edge effect and sensitivity to global mismatch. The only drawback of the original scheme is the fact that it needs optimization.

This averaging effect can be seen as a moving average around the input signal (figure 5.8). So, gradual fluctuations are improved moderately whereas abrupt fluctuations are improved dramatically. Averaging acts like a low-pass filter in the frequency domain (that is why the DNL improves the most).

Several optimization schemes for the design of the averaging network have been proposed [Bul 97b][Abi 01]. [Bul 97b] focuses on the optimum value of the averaging resistance when the load resistance is infinite. [Abi 01] focuses more on a general optimization where the averaging principle is studied as it was a matched filter design. The averaging network is treated as a filter with an impulse response and so on.

The averaging advantage is dependent on the ratio of the load resistance and the averaging resistance. A smaller averaging resistance gives a better averaging effect but on the other hand decreases the gain of the preamplifier. In [Bul 97b] the optimized result was a 3-fold decrease of the DNL values. Also only the amplifiers which are in the linear region can take part in the averaging. So, also the gate-overdrive voltage plays an important role.

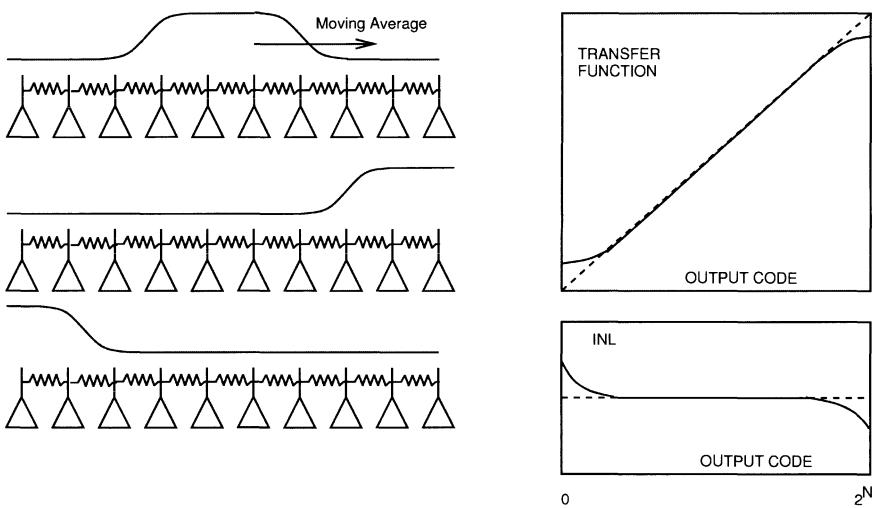


Figure 5.8: Bending Problem in Averaging ADC. Effect on transfer function and INL.

The main disadvantage of the averaging technique is the bending of the output curve, as on both sides of the input range the string of resistors stops (see figure 5.8). A solution could be to sacrifice a part of the input range to keep a linear behaviour. This is done when the design uses dummies to end the left and right part of the resistor string (in fact these dummy preamplifiers simulate the behavior at the end

of the resistor string). But as the trend for digital CMOS processes is to further reduce the supply voltage, introducing overrange sacrifices the signal swing even more. If a circuit could simulate an infinitely long row of amplifiers and averaging resistors, the original signal swing could be preserved. This is done in two designs where resistors with a special value are inserted at the edges to simulate an infinitely long array of resistors [Schol 02].

5.3 ARCHITECTURE OF 8-BIT CONVERTER

As discussed in the previous section, the use of analog preprocessing techniques is necessary to reduce the power consumption of medium-resolution analog to digital converters. Here two of the three analog preprocessing techniques are used to implement the 8-bit converter. The implemented interpolating/averaging architecture is shown in figure 5.9.

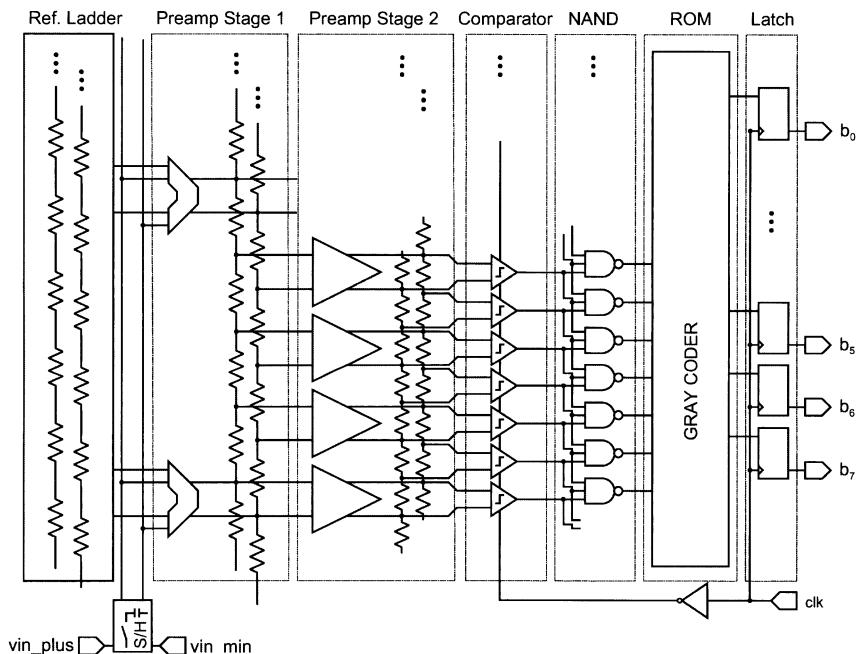


Figure 5.9: Architecture of the 8-bit Interpolating/Averaging Converter.

The architecture is based upon the flash ADCs. The analog signal processing is fully differential to have an improved dynamic performance. The front-end sample and hold samples the input signal. This sampled input value is compared with the reference ladder (differential) and the result is amplified in the first amplifier array

(figure 5.10). To reduce the input-referred capacitance this preamplifier array is interpolated $nr_{INT,st1}$ times.

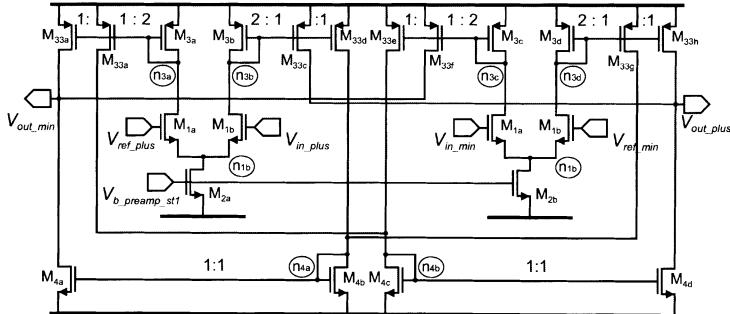


Figure 5.10: Architecture of the first amplifier stage of the Interpolating/Averaging Converter.

To have enough gain to reduce the offset of the comparator, this first amplification stage is followed by a second amplification stage. This second amplification stage is also an interpolated array of amplifiers (interpolated $nr_{INT,st2}$ times). The second preamplifier schematic is shown in figure 5.11.

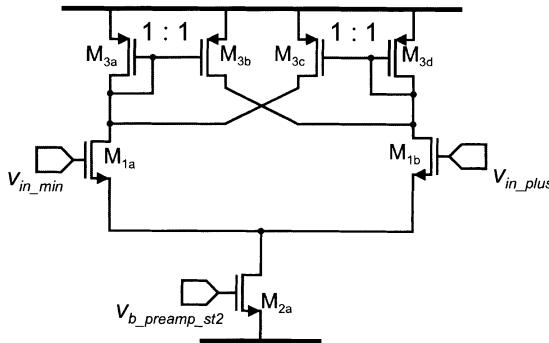


Figure 5.11: Schematic of the second preamplifier.

Both preamplifier stages use averaging to improve static performance [Schol 02]. The outputs of the second preamplifier stage steer the regenerative comparators. A digital back-end, shown in figure 5.12, performs additional error correction (e.g. against bubble errors) and encodes the thermometer coder output from the comparators in Gray code, which is synchronized at the output by a latch. This digital back-end is identical as the one used in flash converters. So the speed requirement should easily be met.

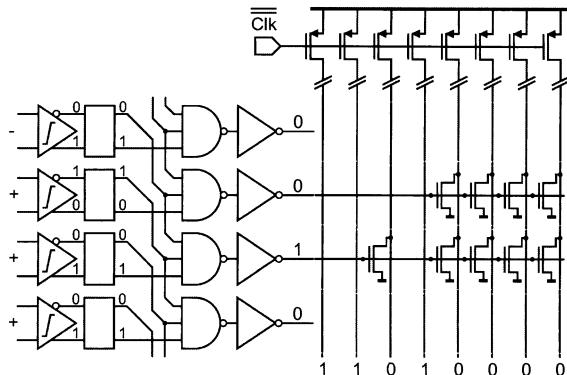


Figure 5.12: Digital Correction and ROM Schematic.

5.4 SYSTEMATIC DESIGN OF THE A/D CONVERTER

To obtain a good performance and a low power budget, a systematic high-level optimization procedure has been implemented. The high-level design equations will be deduced in this section and will be used in the optimization phase to optimize the performance of the 8-bit converter.

5.4.1 Design phase

The specifications that are derived during the system-level specification phase are the input to the converter design phase. The design of the converter is performed hierarchically. The numerical results have been obtained by using a $0.35 \mu\text{m}$ CMOS technology and assuming 8 bits of resolution and a sampling speed of 200 MHz. First, some architectural decisions have to be made. Both static and dynamic performance are taken into account, resulting in specifications for mismatch and admissible phase shift for the different building blocks. These then serve as input to design these blocks at the circuit level.

5.4.1.1 Architectural-level sizing

Considering that the offset voltages of all the comparators in a full flash architecture are independent variables with a normal distribution, then a Monte-Carlo simulation can be used to estimate the design yield as a function of the total equivalent input-referred offset. For these simulations a targeted INL of 1.0 LSB and a targeted DNL of 0.5 LSB were used. Using averaging techniques, the DNL can be improved by a factor of nr_{AVG} , while the INL can be improved by $\sqrt{nr_{AVG}}$ [Bul 97b]. Taking this into account, Monte-Carlo simulations resulted in the plots

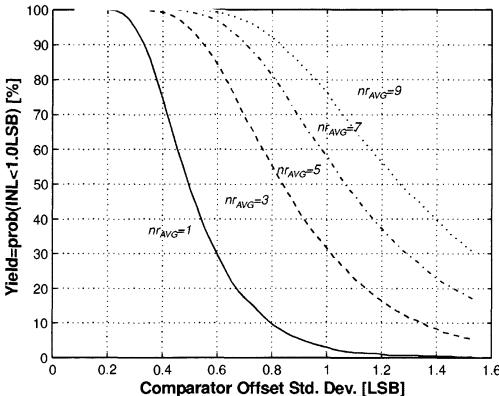
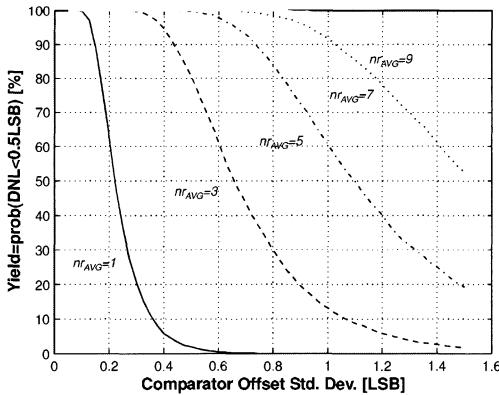


Figure 5.13: Estimated yield as a function of the total input referred offset with the amount of averaging nr_{AVG} as a parameter.

shown in Fig. 5.13: the yield is plotted as a function of the offset, with the amount of averaging nr_{AVG} as a parameter varying from 1 (i.e. no averaging) to 9.

This dependence on the amount of averaging nr_{AVG} is implemented in a lookup table to speed up the circuit-level optimization later on. With e.g. an averaging of 9 ($nr_{AVG} = 9$) the simulations yield a constraint for the admissible total equivalent input-referred offset:

$$\sigma_{total, offset} \leq 0.7 \text{ LSB} \quad (5.1)$$

From Fig. 5.9, the total equivalent input-referred offset $\sigma_{total, offset}$ can be calculated as:

$$\begin{aligned}\sigma_{total, offset}^2 &= \sigma_{preamp_st1, offset}^2 + \left(\frac{\sigma_{preamp_st2, offset}}{A_{preamp_st1}} \right)^2 \\ &\quad + \left(\frac{\sigma_{comp, offset}}{A_{preamp_st1} \cdot A_{preamp_st2}} \right)^2\end{aligned}\quad (5.2)$$

where $\sigma_{preamp_st2, offset}$ is the input-referred offset of the preamplifier stage 1, $\sigma_{preamp_st2, offset}$ is the input-referred offset of the preamplifier stage 2, and $\sigma_{comp, offset}$ is the input-referred offset of the comparator stage. The latter term is negligible if the gain in the preamplifiers is high enough. From statistical behavioral modeling [Vdpl 99] and technological constraints of the process used ($0.35 \mu\text{m}$ CMOS), it can be calculated that a total gain of 15 is sufficient for the comparator to have negligible contribution in the total equivalent input-referred offset. Hence the constraint becomes:

$$\begin{aligned}A_{preamp} &= A_{preamp_st1} \cdot A_{preamp_st2} \\ &= f(INL, technology) \geq 15\end{aligned}\quad (5.3)$$

In our design A_{preamp} was chosen to be 20. Thus mismatch and speed no longer have to be traded off for the comparator, allowing to optimize the comparator for speed.

Apart from the mismatch constraint and the comparator speed, the admissible phase shift for the preamplifier is also determined in this stage of the design. This is a consequence of the variable input slope a particular comparator sees at its input. This variable slope gives rise to a variable delay through the preamplifier and generates harmonic distortion. Fig. 5.14 indicates the problem. The input signal at the input of the deciding comparator can be a steep slope or a slow slope. Depending on this slope and the fact that the preamplifier can be modelled as a first-order RC circuit gives rise to the harmonic distortion.

In [Vdpl] a equation has been derived for the resulting third-order distortion as a function of the internal poles of the preamplifier stages. Although the given equations were derived for bipolar preamplifiers, similar equations can be derived for CMOS amplifier stages [Vand 02]:

$$HD_3 \approx \frac{2g}{3\pi} \frac{f_{in}}{f_b}, \text{ where } g \approx e^{-2b_n \frac{(V_{GS} - V_T)f_{in}}{V_{fs} f_b} - 1} \quad (5.4)$$

V_{fs} is the full-scale input range, f_{in} is the input frequency and f_b is the bandwidth of the preamplifier, g represents the normalized delay $\delta t_d / BW$ of the preamp, b_n is the relative output level. The normalized delay is worst-case around the mid-codes i.e. when $b_n = 0.5$. The results of the above equations are depicted in Fig. 5.15: for

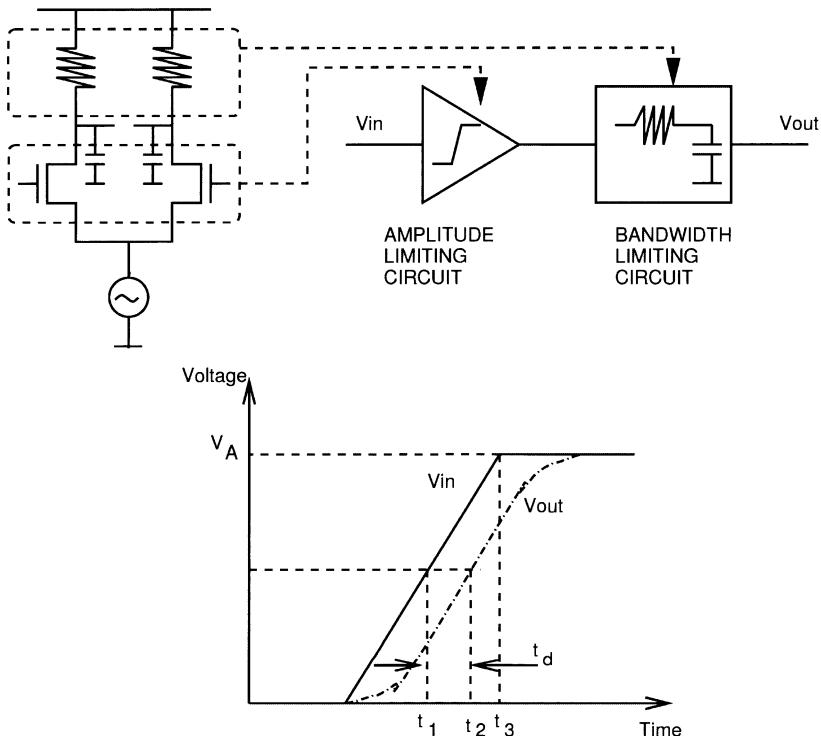


Figure 5.14: Explanation of the problem in ADC due to the pole of the preamplifier.

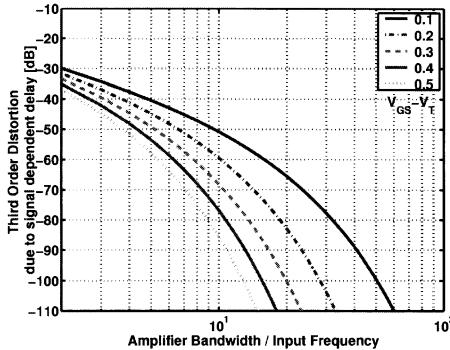


Figure 5.15: Resulting Third Order Distortion as a function of the preamplifier bandwidth/input frequency ratio for a full scale input of 1.25 V.

our example the targeted 50 dB distortion would result in a constraint of 10°phase shift at Nyquist frequency for a $V_{GS} - V_T$ of 0.3 V:

$$\theta_{Nyquist} \leq a \tan\left(\frac{1}{6}\right) \approx 10^\circ \quad (5.5)$$

5.4.1.2 Circuit-level sizing

The architectural-level design resulted in constraints in terms of gain ($A_{preamp} > 15$) and bandwidth of the preamps (e.g. $\theta_{Nyquist} \leq a \tan\left(\frac{1}{6}\right) \approx 10^\circ$), and admissible input-referred offset (e.g. $\sigma_{total, offset} \leq 0.7 \text{ LSB}$) for the different building blocks. Using these constraints, each of the building blocks can be sized as will be discussed in detail in the following paragraphs for each block: SHA, fully differential ladder, 1st-stage preamplifier, 2nd-stage preamplifier, comparator and digital back-end.

Sample & Hold

The differential SHA was based on the architecture presented in [Bro 97] using the clock-boosting technique. The schematic is shown in figure 5.16. Three modifications were done: (1) the gate was boosted with a fixed voltage, (2) special attention was paid to the clock recovery and timing, and (3) a PMOS transistor was added in parallel with the NMOS switch transistor.

The SHA was designed to steer a load of 5 pF (worst-case estimate of the total input capacitance of the preamplifiers based on mismatch constraints) with an input swing of 0.8V (the input swing is chosen by the designer and fixed during the optimization of the preprocessing chain later on). The simulated 3rd harmonic is 68dB and the 5th harmonic is 83dB at a sampling rate of 200 MS/s.

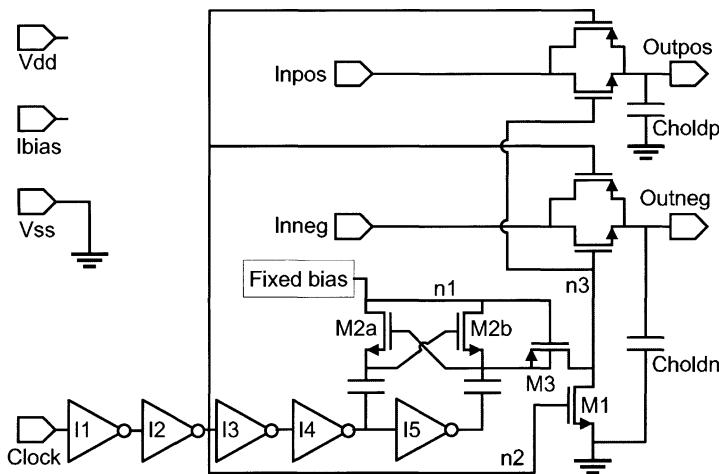


Figure 5.16: Schematic of Sample & Hold.

Reference ladder network

The reference ladder has to be properly sized in order to avoid feedthrough. A first-order estimation of the feedthrough to the midpoint (which is the worst case) of the reference ladder is given by [Ven 96]:

$$V_{mid}/V_{in} = \frac{\pi}{4} fin R_{ladder} C \quad (5.6)$$

In this equation f_{in} is the input frequency, R_{ladder} is the total resistance in the case of one ladder and C stands for the total coupling capacitance from the input to the reference ladder (the gate-source capacitance of the input transistors of the preamplifiers). With this equation, the maximum resistance R_{ladder} is calculated to be 14Ω .

Preamplifier stage 1

The schematic of the first-stage preamplifier is shown in Fig.5.10 . The input-referred offset was calculated using the ISAAC tool [Gie 89]:

$$\sigma_{preamp_st1}^2 = 4 \left[\sigma_{M_1}^2 + \left(\frac{\sigma_{M_{33}}^2}{\sqrt{2}} \right) + \sigma_{M_3}^2 + \sigma_{M_4}^2 \right] \quad (5.7)$$

The gain A_{preamp_st1} of this preamplifier is a function of the number of averaging nr_{AVG} and the number of interpolations nr_{INT} . In [Bul 97a] a new preamplifier

topology was proposed that has a high impedance load, which is beneficial for averaging. The preamplifier used in our design exhibits the same advantage of high intrinsic impedance load (formed by transistors M₃₃ and M₄). Thus the gain in the preamplifier and the averaging effect no longer have to be traded off [Bul 97a].

Using macro models for the amplifiers, a closed-form expression for the overall gain of the preamplifier for a certain number of averaging nr_{AVG} and number of interpolations nr_{INT} , was calculated using the ISAAC tool [Gie 89]. Similar macro models were used to derive equations for other values of nr_{INT} and nr_{AVG} . Comparing the different equations finally resulted in a closed-form expression for the gain:

$$A_{preamp_st1} = -2 \cdot \frac{gm_{m1}}{g_{AVG}} \cdot \frac{nr_{INT} (nr_{AVG} + 1)^2}{2^3} \quad (5.8)$$

This expression is a function of the amount of averaging nr_{AVG} , the number of interpolations nr_{INT} and g_{AVG} is the averaging resistor conductance.

Not only the gain, but also the frequency behavior is affected by the averaging. Expressions were derived for the dominant pole as a function of both the number of averaging nr_{AVG} and the number of interpolations nr_{INT} :

$$f_{dominant_st1} = \frac{1}{2\pi \cdot f(nr_{AVG}, nr_{INT}) \cdot R_{AVG} \cdot C_{load}} \quad (5.9)$$

where $f(nr_{AVG}, nr_{INT})$ is a fit factor extracted from simulations (see Table 5.1). This fit factor $f(\cdot)$ is a function of both the number of averaging nr_{AVG} and the number of interpolations nr_{INT} .

Fit factor f	3	5	7
nr_{INT}	2	2.30e-2	1.04e-2
	4	3.53e-3	1.60e-3

Table 5.1: Fit Factor for dominant pole of preamplifier for a certain combination of the number of interpolations and the amount of averaging.

Preamplifier stage 2

The second-stage preamplifier is depicted in Fig. 5.11. The mismatch contribution is given by:

$$\sigma_{in_st2}^2 = \sigma_{m1}^2 + 2\sigma_{m3}^2 \left(\frac{gm_{m3}}{gm_{m1}} \right)^2 \quad (5.10)$$

As was the case for the first preamplifier, also this 2nd-stage preamplifier has a high output impedance. Equation 5.8 also gives the gain of the 2nd-stage preamplifier

but then as a function of the amount of averaging $nr_{AVG,st2}$ and the number of interpolations $nr_{INT,st2}$.

Also, for the frequency behaviour the same approach as for the 1st stage preamplifier is followed. An equation for the dominant pole, similar to equation (5.9), has been derived.

Regenerative comparator

The comparator used in this A/D converter is a very fast regenerative structure based on cross coupling two inverters. More information about this comparator can be found in [Uyt 00]. The comparator is shown in figure 5.17.

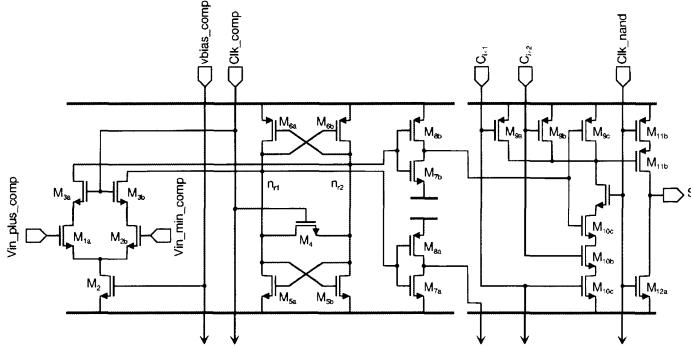


Figure 5.17: Schematic of regenerative comparator.

During the reset phase the voltages on the regenerative nodes n_{r1} and n_{r2} can be calculated as follows [Uyt 00]:

$$(v_{n_{r1}} - v_{n_{r2}}) = (v_{n_{r1}} - v_{n_{r2}}) (t_0) \cdot \exp\left(\frac{g_{eq,res}}{C_{eq,res}} \cdot t\right) + v_{in_comp} \cdot \left(1 - \exp\left(\frac{g_{eq,res}}{C_{eq,res}} \cdot t\right)\right) \quad (5.11)$$

where:

$$g_{eq,res} = g_{m5} + g_{m6} - g_{o5} - g_{o6} - 2g_{ds4} - g_{ds3}, \quad (5.12)$$

and $C_{eq,res}$ equals the total capacitance on the regenerative nodes n_{r1} and n_{r2} . t_0 is the initial start point of reset. The last term in this expression is negligible, as immediately after the clock signal goes down, transistor M_3 still works as a cascode transistor resulting in a high impedance. As soon as transistor M_3 is switched off completely, the impedance becomes infinite. Therefore, the reset time constant can be approximated by:

$$\tau_{res} \cong \frac{C_{eq,res}}{g_{m5} + g_{m6} - g_{o5} - g_{o6} - 2g_{ds4}} \quad (5.13)$$

In order to have reset, $2g_{ds4} + g_{o5} + g_{o6} > g_{m5} + g_{m6}$. On the other hand, the equivalent resistance $1/(2g_{ds4} + g_{o5} + g_{o6} - g_{m5} - g_{m6})$ should be high enough to cause a relatively large initial voltage imbalance in the two regeneration nodes n_{r1} and n_{r2} .

During the regeneration phase (clock is low) the injection of the current imbalance stops, and the conductance of the switch M₄ drops to zero. The regeneration speed is governed by a positive pole, approximately given by:

$$p_{reg} \cong \frac{g_{m5} + g_{m6} - g_{o5} - g_{o6}}{C_{eq,reg}} \quad (5.14)$$

where $C_{eq,reg}$ is the total capacitance on the regenerative nodes n_{r1} and n_{r2} .

Digital back-end logic

The outputs of the comparators form a thermometer code: all comparator outputs below the input level are ‘1’ and vice versa.

Under such conditions, the thermometer code can easily be converted in a binary code by transforming this code in a 1-of-n code followed by a ROM with a binary pattern. A flash converter with this type of structure typically suffers from two problems: bubbles in the thermometer code and metastability [Por 96]. Very fast input signals (near Nyquist frequency) can cause a situation where a ‘1’ is found above a ‘0’. The simplest circuit that can detect this is a 3-input and-gate to ensure that only a single one drives the ROM, as depicted in Fig. 5.9.

Metastability occurs when the applied input signal-difference is very small and the comparator is completely balanced for a short period of time. Therefore, the logic gates driven by that comparator output might interpret the output from the comparator wrongly. As a consequence, zero, one or two ROM lines might be selected leading to severe errors in the digital output code. Therefore, in this design, it was made sure that there is always a valid logic level at the output of the comparator to drive the Gray-encoded ROM. To achieve this, one can use the circuit presented in [Por 96]. Another possibility to accomplish the same effect is using two asymmetric inverters (i.e. the toggle point of the inverter is shifted). This is done by the inverter M7/M8 in Fig. 5.17.

Sizing plan

Combining these equations with the set of constraints resulting from architectural-level synthesis, a full design plan for the converter was derived. The architectural design resulted in three constraints for the design of the preamplifier stages:

$$A_{\text{preamp_st1}} = 10, \quad A_{\text{preamp_st2}} = 2 \quad (5.15)$$

$$\theta_{Nyquist} \leq \alpha \tan\left(\frac{1}{6}\right) \approx 10^\circ \quad (5.16)$$

$$\sigma_{\text{preamp_st1}}^2 \leq \frac{3}{4} (0.7 \text{ LSB})^2, \quad \sigma_{\text{preamp_st2}}^2 \leq \frac{1}{4} (0.7 \text{ LSB})^2 \quad (5.17)$$

With these constraints, and the complete set of design equations derived, all transistors can be sized. As the inter-dependency of the different design variables is high, the sizing plan has been formulated as one (global) constrained optimization problem that has been resolved using Advanced Simulated Annealing and where the power consumption and chip area of the circuit are minimized. The phase shift constraint is evaluated using equation (5.16) during optimization. The offset constraint is implemented as a lookup table and checked as the amount of averaging nr_{AVG} evolves during optimization. The overdrive voltages $V_{GS}-V_T$ of the preamplifiers, the lengths L of the transistors, the biasing currents and the averaging resistor values r_{AVG} are input variables of the optimization. A gate overdrive voltage $V_{GS}-V_T$ of 0.3 V and 0.2 V was chosen as starting point for the 1st, respectively 2nd stage preamplifier. The input range was fixed during optimization as was the number of interpolations, which was chosen $nr_{INT,st1}=4$ and $nr_{INT,st2}=2$.

The optimization of this converter has been done with a simulated annealing algorithm. The optimization problem is formulated as:

$$\underset{\underline{x}}{\text{minimize}} \sum_{i=1}^k w_i \cdot f_i(\underline{x}) \quad \text{such that} \quad \underline{g}(\underline{x}) \leq 0 \quad (5.18)$$

where \underline{x} is the set of independent optimization variables, $\underline{f}(\underline{x})$ is a set of k objective functions, and $\underline{g}(\underline{x})$ denotes a set of l constraints and where w_i are weighting coefficients. Constraint functions are equationed such that a constraint is satisfied when $\underline{g}(\underline{x}) \leq 0$. Being an unconstrained technique, simulated annealing manipulates a single scalar function called *cost function*, $C(\underline{x})$. The constrained optimization in equation (5.18) can be transformed in an unconstrained optimization problem as follows:

$$\underset{\underline{x}}{\text{minimize}} C(\underline{x}) = \sum_{i=1}^k w_i \cdot f_i(\underline{x}) + \sum_{j=1}^l w_j \cdot g_j(\underline{x}) \quad (5.19)$$

The cost function is built by a weighted sum of functions that force the optimization to evolve to *operational* (saturation/linear region), *functional* (design requirements fulfilled) and *applicable solutions* (specifications met). Within this

last design subspace, trade-offs are optimized to result in a solution with minimal area and power [16]. These four categories of cost terms have weighting terms which typically differ an order of magnitude in order to guide the optimization. First the optimization space is scanned for operationally correct spaces, then the circuit needs to be functionally working, then specifications need to be fulfilled and finally area and power consumptions are minimized.

For the MOS transistors the input set is for optimization chosen to be:

$$\underline{x}_{Mi} = \{L, V_{GS} - V_T, I_{DS}\} \quad (5.20)$$

i.e. parameters that control the operating point of the transistors.

A level-1 Spice model was encapsulated in a separate Matlab routine to calculate the device characteristics during optimization. Given L , $V_{GS} - V_T$ and I_{DS} , the routine returns W , g_m , g_o , σ and the parasitic C_{GS} , C_{GD} . Other device-level evaluations (e.g. BSIM) could be used as well, by either a Matlab routine or an external C-code routine which can be invoked easily from within the Matlab environment.

For designing the averaging resistors, an additional routine was added in Matlab. The routine takes the resistive value R_{AVG} and the length L_{AVG} as inputs and decides in which layer to implement the resistor (e.g high-resistive poly, low-resistive poly) such as to minimize parasitic capacitance. The routine returns the width W_{AVG} and the parasitic capacitance C_{AVG} .

The amount of averaging nr_{AVG} is needed to calculate the averaging effect and thus the exact constraint on the admissible input-referred offset $\sigma_{total,offset}$. The amount of averaging nr_{AVG} is however not an independent variable, and is calculated from the $V_{GS} - V_T$, the gain and the linear output range of the preamplifier stages. This loop is resolved by choosing $nr_{AVG,st1}$ and $nr_{AVG,st2}$ as input variables and forcing them to be equal to the derived actual nr_{AVG} .

The simulated annealing loop, executing the overall circuit-level optimization, was implemented in the Matlab environment and uses the Advanced Simulated Annealing (ASA) C-routine as actual algorithm. About 30 trials were needed to obtain the final result. One trial takes 14 min on a Sunblade 1000 workstation.

The sizing of the SHA, the reference ladder network and the digital back-end was not included in the global optimization. The SHA was designed using the ELDO-simulator within an optimization loop, while the digital back-end was designed manually.

5.5 LAYOUT OF THE CONVERTER

As the specifications push the designs closer to the technological boundaries, chip design has become layout driven and parasitics have to be taken into account during design.

The floorplan follows directly from the block diagram in Fig.5.1. The result is depicted in Fig. 5.18. The SHA was inserted on the top. From left to right, the differential ladder network, the 1st and 2nd-stage preamplifiers, the comparators and the digital back-end are placed.

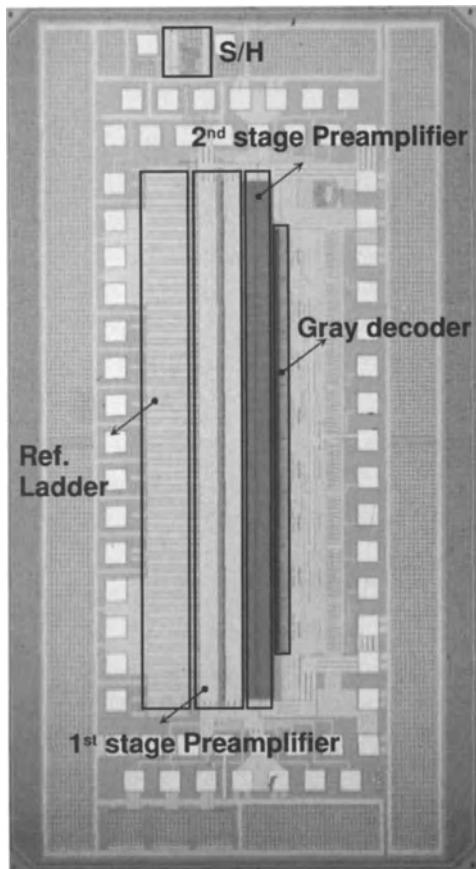


Figure 5.18: Layout of the 8-bit Converter.

Analog and digital power supplies have been separated to avoid cross-coupling from the analog to the digital part. Around the perimeter of the chip 1 nF of decoupling capacitance has been integrated to provide stable power supplies.

The reference ladder was implemented in metal-1 layer. Dummies were added to provide identical surroundings. An additional decoupling capacitance of 10x30pF was added to each ladder to provide stable reference levels.

The layout of the preamplifiers and the routing was done manually; devices were generated using the LAYLA tool [Lam 95], placement of the different modules (1st & 2nd-stage preamplifier) was done using the MONDRIAAN tool [Gie 98]. Internally an additional 500 pF of decoupling capacitance was added. Guard rings were used to reduce substrate (digital) noise coupling. A routing channel has been inserted between the 1st and 2nd-stage preamplifiers. Although this kind of task is automated in digital layout, in analog layout this is still a manual job, as equal delay is important in these connections.

The layout of the digital back-end was done combining Virtuoso from Cadence and the MONDRIAAN tool. The layout of the comparator was done manually as was the internal routing, transistors were generated using the device generator from the LAYLA tool. To increase the speed, the ROM output capacitance has to be minimized. Therefore, the drain area has been made equal to the area of one via (minimum area) as depicted in Fig. 5.19. The ROM cell is handcrafted, 8 ROM cell constitute a ROM line as depicted in Fig. 5.19. The MONDRIAAN tool is used to place the ROM cell and connect the cells using a listing of the Gray code as input. Generation of the ROM is done within 1 minute.

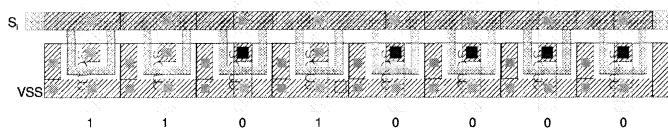


Figure 5.19: Layout of a ROM line.

The clock distribution is critical for mixed-signal designs, and available digital tools cannot deal with the specific analog requirements. A buffered binary clock tree takes care of equal delay, which would otherwise deteriorate the dynamic performance. The design and layout of this clock buffer was done manually as shown schematically in figure 5.20.

5.6 EXPERIMENTAL RESULTS

The A/D converter has been designed for the specifications listed in Table 5.2. The A/D converter was processed in a 0.35 μ m CMOS process. See the chip micro-

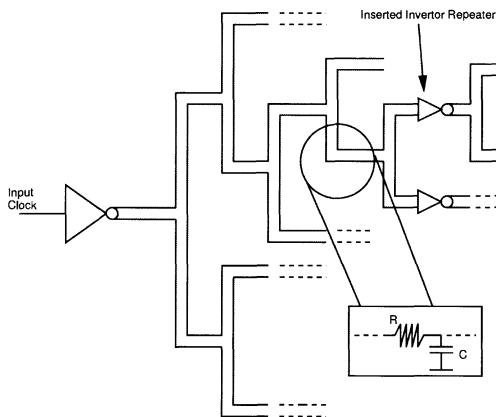


Figure 5.20: Clock buffer.

photograph in Fig.5.18. All biasing was generated on the substrate used to measure the chip to minimize in-coupling noise. The analog preprocessing chain runs from a 3 V power supply, the digital back-end runs at 2.5 V. All measurements were done at full speed of 200 MS/s. The analog preprocessing chain consumes 285 mW, the reference ladder consumes 250 mW and the digital part consumes 120 mW worst case.

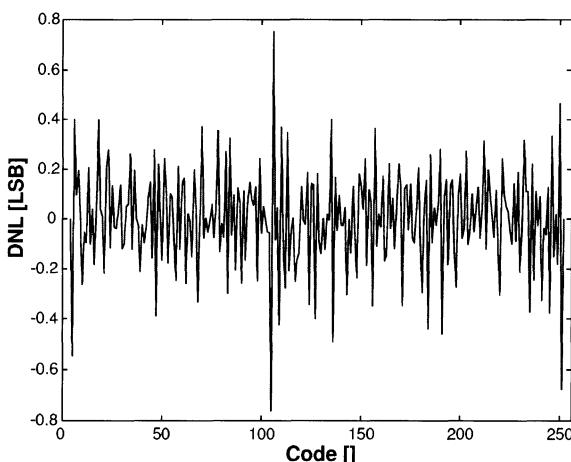


Figure 5.21: Measured DNL of the 8-bit converter.

The measured static performance is shown in Fig.5.21: an INL ≤ 0.95 LSB and a DNL ≤ 0.8 LSB were measured. The dynamic performance is shown in Fig.5.22. A Signal-to-Noise-and-Distortion-Ratio (SNDR) of 44.3 dB is achieved at low frequencies; at 30 MHz a SNDR figure of 43 dB was measured.

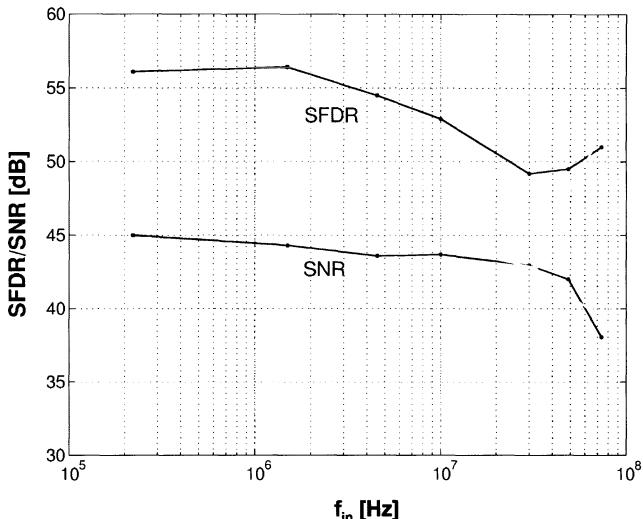


Figure 5.22: Measured dynamic performance of the 8-bit converter.

The measured performance is summarized and compared to the specified values in Table 5.2. All results are comparable to what had been predicted during the sizing.

5.7 CONCLUSIONS

The systematic design of an 8-bit interpolating/averaging 200 MS/s Nyquist-rate A/D converter has been presented. Using behavioral models the system specifications are translated in offset and phase shift constraints that steer the global optimization of the converter at the circuit level. The chip was processed in a standard 0.35 μ m CMOS process. Measurements on the processed chip yielded good results: a DNL/INL figure of 0.8/0.9 LSB has been obtained. At an input frequency of 30 MHz and at full clock speed a SNDR ratio of 43 dB was measured. These results are comparable to the simulated values. The presented systematic design compares favorable to an earlier manual design. In total the manual design took about 6 months, the newly presented approach resulted in functional silicon within 2.5 months, which is a significant speed-up in design time.

	Specification	Unit	Target value	Measured Design
Static	Resolution N	# bits	8	8
	INL/DNL	LSB	< 0.5 LSB	0.8/0.9
	Parametric Yield	%	99.9	-
Dynamic	SFDR	dB	> 45	46
	SNR	dB	> 40	44.3dB@1.5MHz 42.7dB@40MHz
	Sample frequency	MS/s	200	200
Environmental	Conversion rate	-	1 code/clock cycle	1 code/clock cycle
	Input capacitance	pF	<5	4.8
	Input range	V ptp	>0.5	1.3 V
	Latency	-	-	1 clock cycle
	Output load	pF	10	-
	Power supply	V	3.3	3.3/2.5
	Digital levels	-	CMOS	CMOS
	Coding	-	Gray code	Gray code
Optimization	Power	mW	min	655
	Area	$\mu\text{ m}^2$	min	1400x2400

Table 5.2: Performance summary of the 8-bit converter.

References

- [Abi 01] M. Choi and A. Abidi, “A 6-Bit 1.3-gsample/s Flash ADC in 0.35 um CMOS”, *IEEE Journal of Solid-State Circuits*, vol. SC-36, no. 12, pp. 1847–1858, December 2001.
- [Bro 97] T. Brooks and al., “A 16b Sigma-Delta Pipeline ADC with 2.5 mhz Output Data Rate”, in *Proceedings International Solid-State Circuits Conference*, Feb. 1997, pp. 404–405.
- [Bul 97a] K. Bult and A. Buchwald, “An embedded 240-mw 10-bit 50MS/s CMOS ADC in 1 mm²”, *IEEE Journal of Solid-State Circuits*, vol. 21, no. 12, pp. 1887–1895, Dec. 1997.
- [Bul 97b] K. Bult and A. Buchwald, “An embedded 240mW 10b 50MS/s CMOS ADC in 1mm²”, *IEEE Journal of Solid-State Circuits*, vol. 32, pp. 1887–1895, Dec. 1997.
- [Fly 96] M. Flynn and D. Allstot, “CMOS Folding A/D Converters with current-mode interpolation”, *IEEE Journal of Solid-State Circuits*, vol. 28, no. 4, pp. 1248–1257, Sept. 1996.
- [Gie 89] G. Gielen, H. Walscharts, and W. Sansen, “Isaac: a symbolic simulator for analog integrated circuits”, *IEEE Journal of Solid-State Circuits*, vol. 24, no. 6, December 1989.
- [Gie 98] G. V. der Plas, J. Vandebussche, G. Gielen, and W. Sansen, “Mondriaan: a tool for automated layout synthesis of array-type analog blocks”, *Proceedings Custom Integrated Circuits Conference*, May 1998.
- [Kat 91] K. Kattmann and J. Barrow, “A Technique for reducing differential non-linearity errors in flash A/D converters”, in *Proceedings International Solid-State Circuits Conference*, San Francisco, Feb. 1991, pp. 94–98.
- [Lam 95] K. Lampaert, G. Gielen, and W. Sansen, “A performance driven placement tool for Analog Integrated Circuits”, *IEEE Journal of Solid-State Circuits*, vol. 10, no. 1, July 1995.
- [Nau 95] B. Nauta and A. Venes, “A 70-MS/s 110-mw 8-b CMOS Folding and Interpolating A/D Converter”, *IEEE Journal of Solid-State Circuits*, vol. 31, no. 8, pp. 1132–1140, Aug. 1996.

- [Por 96] C. L. Portmann and T. H. Y. Meng, “Power-Efficient Metastability Error Reduction in CMOS flash A/D converters”, *IEEE Journal of Solid-State Circuits*, vol. 31, no. 8, pp. 1132–1140, Aug. 1996.
- [Raz 95] B. Razavi, *Principles of Data Conversion System Design*, IEEE Press, 1995.
- [Roo 93] M. Steyaert, R. Roovers, and J. Craninkx, “A 100MHz 8-b CMOS Interpolating A/D Converter”, in *Proceedings Custom Integrated Circuits Conference*, San Francisco, May 1993, pp. 28.1.1–28.1.4.
- [Schol 02] P. Scholtens and M. Vertregt, “A 6b 1.6GS/s CMOS A/D Converter”, in *Proceedings International Solid-State Circuits Conference*, San Francisco, Feb. 2001.
- [Uyt 00] K. Uyttendhove and M. Steyaert, “A 6 Bit CMOS Very High Acquisition Speed Flash ADC with Digital Error Correction”, in *Proceedings Custom Integrated Circuits Conference*, Orlando, May 2000, pp. ?–?
- [Vand 02] J. Vandenbussche, E. Lauwers, K. Uyttendhove, M. Steyaert, and G. Gielen, “Systematic Design of an 8-bit, 200MS/s Averaging Interpolating CMOS ADC”, in *Design and Automation Test Equipment*, Paris, Mar. 2002, pp. 104–108.
- [Vdpl] R. V. de Plassche, *Integrated Analog-to-digital and Digital-to-analog Converters*, Kluwer Academic Publisher, 1994.
- [Vdpl 99] G. V. der Plas and J. V. et al., “Statistical Behavioral Modeling for A/D Converters”, in *Proc. of the IEEE 1999 ICECS*, Sept. 1999, pp. 1713–1716.
- [Ven 96] A. Venes and R. J. V. de Plassche, “An 80MHz 80mW 8b CMOS Folding A/D Converter with Distributed T/H Preprocessing”, in *Proceedings International Solid-State Circuits Conference*, San Francisco, Feb. 1996, pp. 241–243.

Chapter 6

HIGH-SPEED FLASH ADCS

Design issues of a 6 bit, 1 GHz CMOS Flash ADC

Koen Uyttenhove, Michiel Steyaert

Abstract Designing high-speed flash ADCs in a deep submicron technology requires optimized architectures and building blocks. In this chapter, the design of a high-speed 6 bit converter is presented. In the first section, an introduction will be given to flash architectures, after which the design issues of the different building blocks of the converter will be discussed. At the end, the measured results of the implemented converter will be shown.

6.1 INTRODUCTION

High-speed A/D converters used in automated test equipment, oscilloscopes and in digital data reading (for example hard disk drives and digital video disk) often require resolutions as low as 6 bit, sampling speeds as high as possible, and low latency [Alt 01]. Due to the increased levels of integration achieved in the last years, the demand for high speed CMOS analog to digital converters has increased.

Furthermore, $\Delta\Sigma$ converters using multi bit quantizers represent a very interesting alternative to obtain simultaneously high signal bandwidths and high resolutions. These type of converters are in fact complex systems that make use of two extra data converters: a high-speed and low resolution A/D converter, and a high-speed and high accuracy D/A converter.

With this motivation, this chapter discusses the design and measurement of a 6 bit A/D converter implemented in a standard $0.35\text{ }\mu\text{m}$ CMOS technology, that has been especially optimized to reach the highest acquisition frequency.

The outline of this work is as follows. In the next section the converter architecture is shortly discussed. Then the design of the converter building blocks is discussed. In the next section the measurement results are presented. Finally, some final conclusions are drawn.

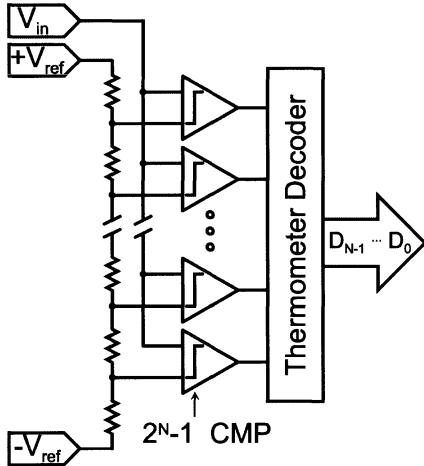


Figure 6.1: Generic Block diagram of a flash A/D converter.

6.2 ADC CONVERTER ARCHITECTURE

As indicated in the first chapter on ADC architectures, flash architectures are typically the simplest and the fastest structures that can be used to implement analog to digital converters. Figure 6.1 presents a block diagram of a N bit flash converter. The resistive ladder subdivides the converter reference voltage ($+V_{ref} - V_{ref}$) in a set of 2^N reference voltages, which are compared in parallel with the analog input signal. A logic decoder converts the thermometer code generated by all the comparators into a binary code that approximates the input signal every clock cycle.

Note that the major advantages (simplicity and parallelism) of flash architectures also presents its main problem: the number of comparators increases exponentially with the resolution specification, leading typically to a large die area and a high power consumption. Normally, this architecture is only used to implement converters with a resolution less than or equal 6 bit.

6.2.1 High-level Offset Demand in Flash ADC

As the most important task of this ADC is the correct comparison between the input level and the reference voltage level, the correct operation of a flash converter depends on the accurate definition of the reference voltages sensed by each comparator. In fact, the offset voltage of each comparator sums directly to each reference voltage.

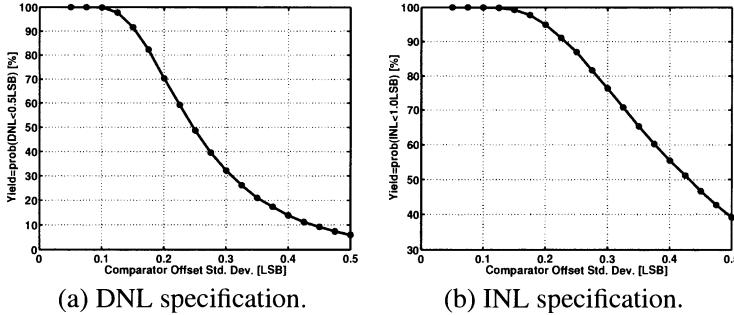


Figure 6.2: Design yield as a function of the standard deviation of the comparator offset voltage (a) for a DNL ≤ 0.5 LSB and (b) for an INL ≤ 1 LSB.

Since the comparator offset voltage is a random variable (which depends on the matching properties of the used technology), it directly influences the differential and integral non linearity characteristics of the ADC. Therefore, the first step in the design of a flash converter consists in deriving an offset voltage standard deviation that guarantees with a high probability that the design complies with a certain performance specification (that is, an offset voltage that ensures a high design yield) [Pel 98].

Consider that the offset voltage of all the comparators are independent variables that follow a normal distribution. A Monte Carlo simulation can be used to estimate the design yield as a function of the offset voltage standard deviation. Also a closed-form expression can be used where the Error Function will appear, but this expression is difficult to obtain for a DNL specification where the difference between two reference voltages is important. The results obtained considering that one wants to comply with a DNL and an INL specification of 0.5 LSB and 1.0 LSB, respectively, are presented in Figure 6.2.

This design was made for a σ_{offset} smaller than 0.12 LSB, to which a yield estimation of about 99 % corresponds.

The next section discusses the design of the converter building blocks.

6.3 BUILDING BLOCKS

The architecture of a flash converter is a very regular structure. Figure 6.3 presents a more detailed block diagram of the flash architecture that will be used in the following description. Note that even the thermometer decoder can be implemented with a simple and regular pattern.

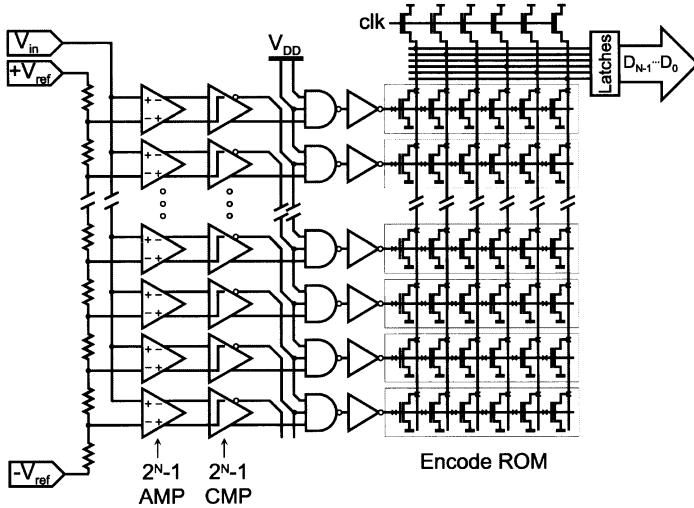


Figure 6.3: Detailed diagram of a flash A/D converter.

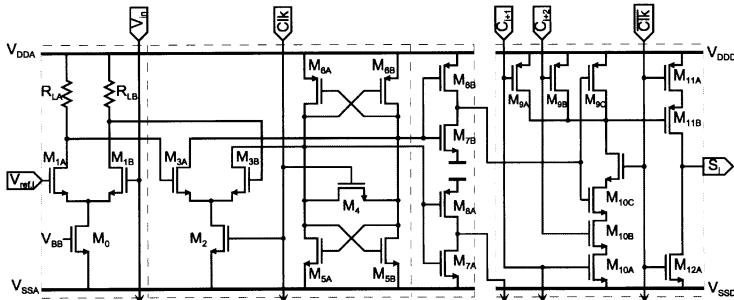


Figure 6.4: Schematic diagram of one segment of the implemented flash A/D converter. Clk and $Clk-bar$ represent the clock signal and the negation of the clock signal, respectively. C_{i+1} and C_{i+2} represent the output of comparators $i + 1$ and $i + 2$ after the invertor gate. S_i represents the select signal for the ROM line i .

Figure 6.4 presents the schematic diagram of segment i of the flash converter. The design of this segment is discussed next, starting with the preamplifier.

6.3.1 Design of the Preamplifier

The amplifier used in front of the comparator is used for two reasons:

- First, it buffers the input signal and the reference lines from the kick back noise caused by the regeneration operation of the comparator.
- Second, since the preamplifier has a gain larger than unity, the input referred offset caused by the comparator circuit is reduced. This is necessary because the typical offset of the comparator can be quite large.

As derived before, the standard deviation of the total input referred offset has to be smaller than 0.12 LSB. To reduce the offset specification, one should increase the size of one LSB as much as possible. Obviously, this implies maximizing the input swing of the preamplifiers.

On the other hand, the preamplifier gain is determined by the available output swing and the overdrive voltage of the differential input pair transistor:

$$A_v = g_m R_L = \frac{I_{BB}}{(V_{GS} - V_T)_1} \cdot \frac{V_{out,swing}}{I_{BB}} = \frac{V_{out,swing}}{(V_{GS} - V_T)_1} \quad (6.1)$$

Consequently, one should increase the input preamplifier input and output swing and use reduced overdrive voltages for the input and bias transistor. In this design the range of the preamplifier input and output voltages are 1.0 to 2.6 V and 2.3 to 3.3 V, respectively. The overdrive voltage of the bias and input transistors is about 0.30 and 0.25 V, respectively. Consequently, the preamplifier gain is about four.

In the next sub-section the design of the comparator is discussed.

6.3.2 Design of Comparator

The comparator used in this flash is a very simple and fast regenerative structure (as this structure has the highest speed possibility in CMOS). During one of the clock phases, the comparator is reset at the same time that the clocked differential pair composed by transistors M_{03A} and M_{03B} injects in the regeneration nodes a current imbalance proportional to the preamplifier output signal. In the next clock phase, the voltage imbalance that exists between the regeneration nodes is amplified by the NMOS and PMOS regeneration loops to digital levels¹.

6.3.2.1 Offset of the comparator

The first aspect to consider in the design of this structure is the offset voltage. Consider the mismatch between each pair of transistors modeled by a random and independent voltage source in series with the gate of one transistor. The variance

¹This means that in one phase the input signal is tracked, and during the clock transition the comparators effectively sample the polarity of the applied signal, avoiding in this way a sample and hold circuit at the front of the converter. Since this architecture does not require a sample and hold circuit, and because the comparators do not have to be linear, one can better understand why these architecture allows to achieve the highest acquisition frequency.

of the input referred offset is then given by the total sum of the offset contribution of the preamplifier and the offset contribution of the comparator (which is divided by the gain of the preamplifier):

$$\sigma_{offset}^2 = \sigma_{M1}^2 + \frac{1}{(g_{m1}R_L)^2} \cdot \sigma_{M3}^2 + \frac{g_{m5}^2}{(g_{m3}g_{m1}R_L)^2} \cdot \sigma_{M5}^2 + \frac{g_{m6}^2}{(g_{m3}g_{m1}R_L)^2} \cdot \sigma_{M6}^2 \quad (6.2)$$

The preamplifier gain ($g_{m1}R_L$) has already been discussed. In this design the size of the transistors M_3 has been reduced relatively to size of transistors M_5 and M_6 . In this way the contribution of M_3 to the capacitance at the regeneration nodes is reduced. However, this has the disadvantage of reducing the transconductance of M_3 relatively to the transconductance of M_5 and M_6 , increasing the contribution of the regeneration transistors to the overall offset voltage. In this design g_{m5} and g_{m6} are about equal, and roughly twice the value of g_{m3} .

The offset voltage of a differential pair can be expressed as a function of the overdrive voltage and the gate area by [Pel 89]

$$\sigma_{Mi}^2 = \frac{A_{VT}^2}{(WL)_i} + \frac{(V_{GS} - V_T)_i^2}{4} \cdot \frac{A_\beta^2}{(WL)_i} \approx \frac{A_{VT}^2}{(WL)_i} \quad (6.3)$$

where A_{VT} and A_β are the threshold voltage and relative current factor mismatch parameters, respectively. The previous approximation is valid for relatively small gate-source overdrive voltages.

Considering the A_{VT} equal to 10 mV μ m, the gate-source areas of the transistors can be calculated.

6.3.2.2 Speed of the comparator

The second aspect to consider in the design of the comparator is the reset and regeneration speed.

During the reset phase the voltage difference between the two regeneration nodes is approximately determined by the following equation [Uyt 00]:

$$(v_a - v_b)(t) = (v_a - v_b)(0) \cdot \exp\left(\frac{g_{eq}}{C_{eq}} \cdot t\right) + g_{m1}R_L \cdot \frac{g_{m3}}{g_{eq}} \cdot v_{in} \cdot \left(1 - \exp\left(\frac{g_{eq}}{C_{eq}} \cdot t\right)\right) \quad (6.4)$$

with the following parameters

$$g_{eq} = g_{m5} + g_{m6} - 2g_{ds4} - g_{ds3} \quad (6.5)$$

$$C_{eq} = C_{dtot5} + C_{gtot5} + C_{dtot6} + C_{gtot6} + C_{dtot4} + C_{dtot3} + C_{gtot7} + C_{gtot8} \quad (6.6)$$

In order to have a reset, one should make $2g_{ds4} + g_{ds3} > g_{m5} + g_{m6}$. On the other hand, one should have an “equivalent resistance” $1/(2g_{ds4} + g_{ds3} - g_{m5} - g_{m6})$ high enough to cause a relatively large initial voltage imbalance in the two regeneration nodes. Because of this trade-off between reset accuracy and initial imbalance, in this design $2g_{ds4} + g_{ds3}$ has been made about twice the value of $g_{m5} + g_{m6}$.

During the regeneration phase the injection of the imbalance current stops, and the conductance of the reset switch drops to zero. The regeneration speed is then governed by a positive pole given by

$$p_{reg} = \frac{g_{m5} + g_{m6}}{C_{dtot5} + C_{gtot5} + C_{dtot6} + C_{gtot6} + C_{gtot7} + C_{gtot8}} \quad (6.7)$$

Note that both the NMOS and PMOS regeneration loops contribute with transconductance and with parasitic capacitance to the definition of this pole. If the size of the PMOS regeneration loop are relatively large compared with the size of the NMOS regeneration loop, the regeneration speed is strongly reduced because of the following reasons. First, the PMOS transistors add too much capacitance to the regeneration nodes and add a limited transconductance. Second, the threshold voltage of the comparator moves towards V_{DD} , leading to a higher overdrive voltage of the NMOS transistors, and consequently to a lower transconductance (for the same current budget).

Consequently, in order to improve the regeneration speed the ratio between the size of the PMOS regeneration transistors and the NMOS regeneration transistors should be smaller than the ratio between the carriers mobility of the NMOS and PMOS transistors (about 3). Furthermore, the current that passes through the current injection transistors M_3 also passes through transistors M_6 . In this way, the transconductance of the PMOS regeneration loop can also be increased. In this design, the transistor sizes have been scaled so that both the transconductance and the capacitances associated with the top and bottom regeneration loops are about equal.

Optimisation of comparator speed

As indicated in previous section, some general design issues can be derived to optimize the speed of the regenerative comparator. In this section, a more mathematical optimisation will be presented.

A simulation of the time constant is done in Matlab, where the drain-bulk capacitances and gate-source capacitances are modelled with the transistor equations. The main matlab code is shown next and is based upon the previous derived equation for the p_{reg} whereby the long channel design equations have been used to simplify the equation. Also an estimation of the ratio of the drain-

bulk capacitance and the gate-source capacitance has been made (by the factor a):

```
% Optimization for speed of two regeneration loops

vt = 0.6;
vdd = 3.3;
vmin = 0.2;
vgst = vmin:0.01:vdd-2*vt-vmin;
kt = 3;
a = 1/3;

% vgsti = 1;

iratio = 0.1:0.1:1.1;
gmc = zeros(length(iratio),length(vgst));
for i = 1:length(iratio),
    gmc(i,:) = (1 + vgst./(vdd-(vgst+2*vt)) * (1+iratio(i)))
        (1 + kt*vgst.^2./(vdd-(vgst+2*vt)).^2 *

end

plot(vgst,gmc);
grid on;
xlabel('Overdrive voltage: Vgst-Vt [V]');
ylabel('Optimization of Comp Speed');
```

Together with this Matlab simulation the regeneration speed of the comparator is plotted in figure 6.5 as a function of the current difference between the NMOS and the PMOS regeneration branch and the gate-overdrive voltage of the NMOS transistor because these two variables are the only independent variables left in this optimization problem.

One can clearly see that there exist an optimum for the comparator speed, this optimum is shifted when the current ratio between the NMOS and the PMOS branch is changed.

The next sub-section addresses the dimensioning of the ladder network.

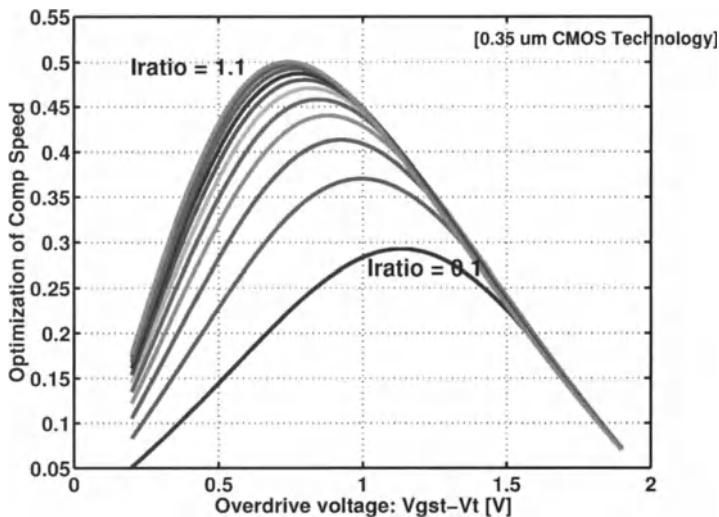


Figure 6.5: Optimisation of comparator speed as a function of current ratio between NMOS and PMOS branch and NMOS overdrive voltage.

6.3.3 Design of the Resistor Ladder

The ladder network subdivides the converter reference voltage in equal reference voltages for each comparator, that should be constant under all conditions. One important source of errors in a flash ADC is caused by the capacitive feedthrough of the input signal to the resistor ladder. For relatively high frequency signals, the converter input signal easily couples to every tap of the ladder through the gate-source capacitances of the differential input pair transistors. Consequently, the voltage at each tap of the ladder network can change substantially from its nominal DC value, degrading in this way the converter performance. Furthermore, two aspects aggravate this effect even more. First, the coupling to a given tap voltage happens not only through the differential pair directly connected to that node, but also through the neighbor differential input pairs [Ven 96]. Second, since the impedance at each node is not constant, and the coupling capacitances are strongly voltage dependant, the coupling to every node is a non linear function of the applied input signal.

6.3.3.1 Model and calculation of ladder feedthrough

To calculate the feedthrough from the input source to the resistive reference ladder. To midpoint of the reference ladder is the worst-case point. Fig. 6.6 presents the

used model to calculate the feedthrough. R is the resistance value for each resistor and C models the feedthrough capacitance from the input source to the ladder, i.e. the gate-source capacitances from the input differential pair.

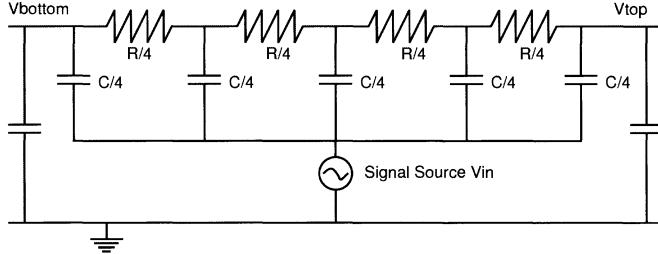


Figure 6.6: Model of Resistive Ladder Feedthrough.

It is assumed that the feedthrough at nodes V_{top} and V_{bottom} is negligible due to proper decoupling. Maximum feedthrough will occur on the mid node: the feedthrough to the middle of the reference ladder is the highest because at this point the equivalent impedance seen when looking into the reference ladder is the highest of all; with a feedthrough from the input voltage with frequency given by:

$$\frac{V_{mid}}{V_{in}} = \frac{\alpha(\alpha + 32)}{\alpha^2 + 32\alpha + 128} \quad (6.8)$$

with $\alpha = \pi f_{in} RC$ If $\alpha \ll 1$ than the equation can be simplified to

$$\frac{V_{mid}}{V_{in}} = \frac{\pi}{4} f_{in} RC \quad (6.9)$$

So, the most effective way of reducing the magnitude of the coupling to a specific node actually consists in reducing the total resistance of the ladder network. Resorting to circuit level simulations it has been found that if the total ladder resistance is smaller than about 100Ω , then the amplitude of coupling to the middle reference tap (the most critical one) is bellow 1 LSB up to an input frequency of 200 MHz. For higher input frequencies, the ladder resistance has to be even more reduced.

In this design, the reference ladder has been implemented usign the first metal layer drawn and using an S shape to obtain a total resistance of about 100Ω .

6.3.3.2 Accuracy Considerations in Monolithic Resistors

Also very important is the matching of the ladder network, since it directly influences the differential and integral non linearity of the converter.

A schematic diagram of a resistor string and its equivalent circuit are shown in Fig.6.7. Actually, the resistor string is folded into a square array in order to save

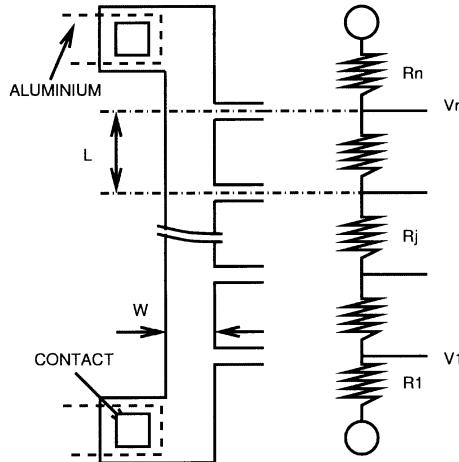


Figure 6.7: Schematic Diagram of resistor string.

chip space and to lessen the converters sensitivity to temperature and process variations. Diffused resistors, ion-implanted resistors, and doped polycrystalline silicon resistors are used in resistor strings in LSI's. Now let us clarify the dependence of variations in R_j on the dimensions of a resistor when variations in resistor width W and length L are independent of each other and probability density conditions are normal. If sheet resistance ρ , is defined as in the relation $R_j = \rho \cdot \frac{L}{W}$ and σ_R is defined as the standard deviation of R_j divided by the average value of R_j , then the following equation can be obtained:

$$\sigma_R = \frac{\sigma_d}{w_0} \sqrt{1 + \frac{1}{D^2}} \quad (6.10)$$

where, W_0 design value of W (average value), L_0 , design value of L (average value), σ_d , standard deviation of W , which is equal to that of L , D dimension ratio of R_j , i.e., L_0/W_0 . As is evident from eq.6.10, σ_R is proportional to $\frac{\sigma_d}{W_0}$ only when $D > 2$. Thus it is advisable to increase W_0 as far as chip size allows, with restriction that $L_0/W_0 > 2$.

With this value, we can calculate the maximum sigma to obtain a N -bit flash converter. The reference voltage, V_{ref} is applied at the endpoints of the resistor string. The absolute values of the tap voltages, V_{tap} , are a fraction, $f(tap)$, of the reference voltage. The accuracy requirement is that any tap voltage should be

accurate within $0.5LSB$ or $\frac{V_{ref}}{2^{n+1}}$ at the n-bit level:

$$V_{tap} = f(tap)V_{ref} \pm \frac{V_{ref}}{2^{n+1}} \quad (6.11)$$

The largest error occurs for the maximum $f(tap)$ which is 0.5 and occurs at the middle of the string. The fractional output is:

$$\frac{V_{tap}}{V_{ref}} = f(tap) \left(1 \pm \frac{1}{2^{n+1}f(tap)} \right) \quad (6.12)$$

The second term in the parenthesis is the error and is $\frac{1}{2^n}$ at the worst case in the middle of the string. In the middle the ratio $\frac{V_{out}}{V_{ref}}$ should be in the following range:

$$\frac{V_{out}}{V_{ref}} = 1/2 \left(1 \pm \frac{1}{2^n} \right) \quad (6.13)$$

This ratio of voltages is the ratio of the resistance of half of the string to the total resistance. Summing the resistors in each term:

$$\frac{V_{out}}{V_{ref}} = \frac{\sum_{i=1}^{i=2^{n-1}} R_j}{R_t} = \frac{\sum_{i=1}^{i=2^{n-1}} R_j}{\sum_{i=1}^{i=2^{n-1}} R_j + \sum_{i=2^{n-1}+1}^{i=2^n} R_j} = 1/2 \left(1 \pm \frac{1}{2^n} \right) \quad (6.14)$$

Assume the resistor values are normally distributed with mean R and standard deviation σ_R . Then the sum of resistors will also be normally distributed with a mean equal to the sums of the means and the standard deviation equal to the square root of the sum of the squares of the standard deviations of the resistors. If any resistor is assumed to be made up of a mean value and a standard deviation then $R_j = R + \sigma_R$ and previous equation can be rewritten as:

$$\frac{V_{out}}{V_{ref}} = \frac{2^{n-1}R \pm 2^{\frac{n-1}{2}}\sigma_R}{2^{n-1}R \pm 2^{\frac{n-1}{2}}\sigma_R + 2^{n-1}R \pm 2^{\frac{n-1}{2}}\sigma_R} = 1/2 \left(1 \pm \frac{1}{2^n} \right) \quad (6.15)$$

The numerator and left term in the denominator are the same term and their sign is taken as positive. To maximize the error effect the negative sign is taken in the right denominator term. With these conventions:

$$1/2 \left(1 \pm \frac{\sigma_R 2^{-\frac{n+1}{2}}}{R} \right) = 1/2 \left(1 \pm \frac{1}{2^n} \right) \quad (6.16)$$

This can be simplified to show that $\frac{\sigma_R}{R}$, the maximum mismatch for less than $0.5LSB$ error is:

$$\frac{\sigma_R}{R} = 2^{\frac{n-1}{2}-n} = \frac{1}{\sqrt{2}\sqrt{2^n}} \quad (6.17)$$

For this prototype, 6 bit resolution is needed so,

$$\left(\frac{\sigma_R}{R} \right)_{6bit} = 0.5\% \quad (6.18)$$

Variations in R_j which depend on inaccuracies in the dimensions of a resistor, e.g., errors in the mask, in mask alignment, in the photo-etching process, etc., and variations in p , are considered to be comparatively random and mutually independent. On the other hand, variations in R_j which depend on the piezoresistance effect, variations in the thickness of the depletion layer, and temperature distribution, are considered to be nonrandom, i.e., they are considered to be deterministic variations.

6.3.3.3 Designing the resistor ladder

The two effects discussed above are now used to design the resistor ladder. Mismatch in resistor string and offset voltage directly contribute to the INL and DNL error of the converter given as below:

$$INL_{MAX} = \frac{V_{ref}}{2} \cdot \frac{\Delta R}{R} + V_{offset} \quad (6.19)$$

If it assumed that maximum positive mismatch occurs in the lower half, maximum negative mismatch occurs in the upper half or vice versa and that i th folder amplifier and comparator contains the maximum offset voltage, then the equation shows the maximum INL for the converter. Hence to reduce the INL, resistor ladder should be maximized. But the minimum ladder impedance for reducing the feed through and preventing the distortion of the input signal from RC effect is given in eq. 6.10.

The above two equations have been used to establish the minimum ladder impedance. In this design for the total input capacitance of about 0.5pf, the total ladder impedance of 100Ω has been used as a tradeoff between minimizing mismatch of resistors, feed through and distortion to the input signal.

The next sub-section concludes the discussion of the design of the building blocks by presenting the design of the converter digital logic.

6.3.4 Design of the Digital Logic

6.3.4.1 Digital logic with no error correction

The outputs of the comparators constitute what is known as a thermometer code: under perfect conditions, all comparator outputs below the input level are ones, and all comparator outputs above the input level are zero. Under such conditions, the thermometer code can easily be converted in a binary code by: first, using a circuit that detects the transition from one to zero; and second, using the one out of 2^N code to address a ROM with the binary code.

A flash converter with this type of structure typically suffers from two problems: first, bubbles (or sparkles) in the thermometer code and metastability [Por 96][Man 90]. These problems are discussed next.

6.3.4.2 Bubble Errors in the ADC thermometer code

For very fast input signals, small timing differences between the response time of the comparators combined with unfavorable offset voltages in the comparators can cause a situation where a one is found above a zero—this is normally called a “bubble” in the thermometer code, because this error resembles bubbles in the mercury of a thermometer.

There exist bubbles from first order and higher-order, referring to the number of zeros situated between two ones in the thermometer code. The case of a first order bubble is shown in figure 6.8(a) and the case of a second order bubble is shown in figure 6.8(b). Indicated on the figure are the delay of the clock between the different comparator stages and the shift of the reference voltages due to offset or other causes.

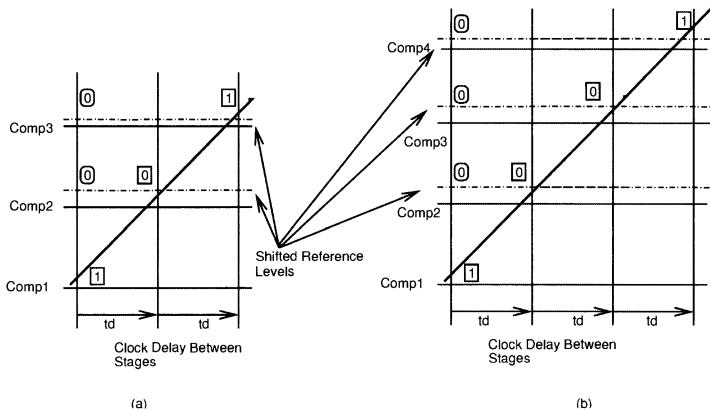


Figure 6.8: Illustration of (a) first order and (b) second order bubbles in the thermometer code.

The probability of encountering a second-order bubble is much smaller than the probability of having a first order bubble in the thermometer code.

The simplest circuit that can detect the transition from a one to a zero is a two input and gate (with the input from the top comparator negated). However, with such a circuit, in the presence of a bubble there are two transitions from one to zero, and consequently two ROM lines are addressed (if the above mentioned digital encoder is used). The output code can then be completely wrong. The output code is then the and-function of all the rom lines which are addressed by the comparators. Even a small error can originate a large glitch in the output code, leading to a serious degradation of the performance of the converter.

6.3.4.3 Improved digital output encoder

The simplest and most often used method to suppressing bubbles consists in using a three input and gate to generate the code that addresses the ROM. With this approach (see Figure 6.3) two zeros have to be found above a one to cause the ROM line to be activated. This simple scheme solves the case of the most typical bubbles in the thermometer code (first order bubbles). Note that if the bubble has two or more zeros it is not corrected with this approach. However, this case is relatively less important for two reasons. First, the case of two zero bubbles is relatively rare since it requires larger timing mismatches between the different comparators. Second, when two ROM lines are selected (because of two zero bubbles), the output code generated in this design is relatively close to the correct value since the ROM has been implemented with a Gray code: the output of the ROM is the logic and between the two selected ROM lines. Since there is only one bit changing between adjacent Gray codes and because the selected lines are close to each other, the resulting code is a good guess of the correct input level.

Several other correction schemes have been suggested by several authors, these correction schemes are able to correct larger errors in the thermometer code (e.g. bubbles with two zeros) but require additional circuitry. These other correction schemes have been summarized here and a small explanation is given beneath.

- The Mangelsdorf Method [Man 90]
- The bit swapping method [Gar 89]
- The Wallace Tree Method [Kae 97]

In the bit swapping method the number of ones is counted in the thermometer code. By swapping the zero-to-one crossings, the number of zero-to-one crossings is reduced leading ultimately to only one zero-to-one crossing (as illustrated in figure 6.9).

This reduces the error induced in the following ROM. The swapping procedure can be optimized as done in the Wallace tree method where the often used addition technique in the digital electronics is used to count the number of ones in the thermometer code. This method is equal to the result obtained with the bit swapping procedure if this method is implemented with maximum order (2^n with n the number of bits). The disadvantage is the large amount of hardware and area necessary to perform this error correction. The Mangelsdorf method consists of taking a majority vote of a comparator and its left and right neighbor.

So in fact, the bit swapping method and the wallace tree method perform a global error correction whereas the mangelsdorf method and the three-input gate method perform a local error correction. As indicated in previous paragraph, the case that

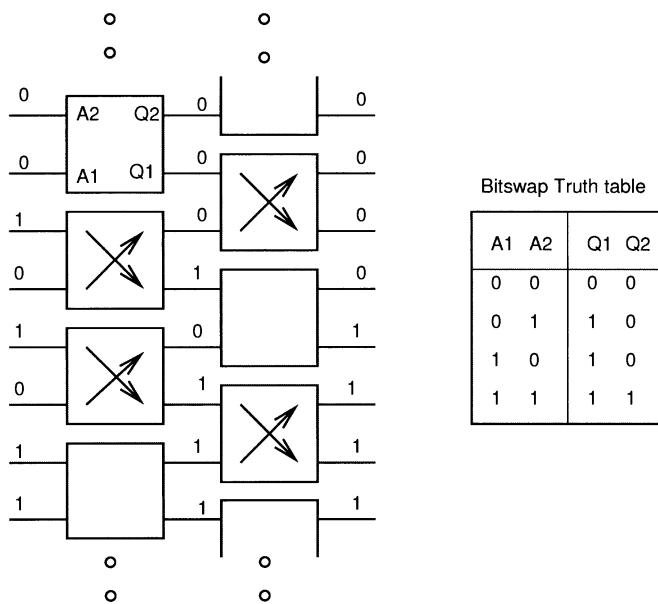


Figure 6.9: Illustration of Bitswap method to remove bubbles in thermometer code.

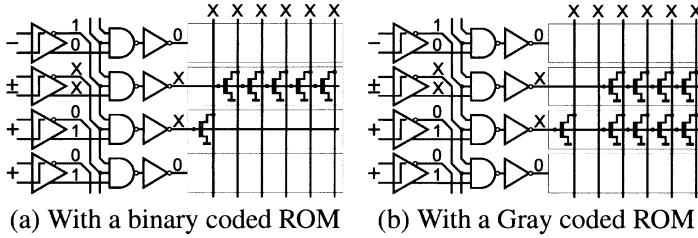


Figure 6.10: Error propagation up to the converter output caused by one undecided comparator. Note that in the case of the Gray coded ROM the error caused by the undecided comparator only originates a completely erroneous code if no line is selected.

larger errors happen is rather low, so in this design the first encoder with the three-input gates is chosen. Especially when going to smaller and smaller technology lengths, the probability of a large delay in the clock circuitry reduces drastically leading to a smaller error probability for the same sampling speed and clock speed.

6.3.4.4 Metastability errors of the comparator

Metastability designates the ability of a comparator to be completely balanced for a short period of time when the applied input signal difference is very small. In a flash converter this means that when the input signal is very close to one of the reference voltages, that comparator might be unable to toggle to a valid logic level. Therefore, the logic gates driven by that comparator output might interpret the input as different levels. As consequence zero, one or two ROM lines might be selected leading to severe errors in the digital output code. Figure 6.10 illustrates a situation caused by metastability where all output bits might be wrong.

At the end of the regeneration phase, the comparator output can be approximated by an expression of the form

$$v_o = v_i \cdot A \exp (+p_{reg} \cdot T/2) \quad (6.20)$$

where v_i is the comparator input voltage difference, A is the comparator input-output gain, T is the clock period, and p_{reg} is the regenerative pole expressed by eqn:regpole. One can consider that the regenerative circuit has an overall gain equal to

$$A_{reg} = A \exp (+p_{reg} \cdot T/2) \quad (6.21)$$

In the presence of a small input signal, a low regeneration gain might be insufficient to amplify the input voltage to a valid logic level, increasing in this way the probabilities of having an error. In short, the metastability error probability (and

the number of errors per unit of time) is inversely proportional to the regenerative gain.

Since the error rates increase exponentially with the acquisition frequencies, the metastability errors might increase by orders of magnitude for very high-speed converters. On the other hand, the error rates are exponentially proportional to the regenerative time constant. Therefore, the most effective way of reducing the metastability error rates is to reduce the regenerative time constant, that is to increase the positive regenerative pole—as described before.

Although this provides an efficient way of reducing the error rates, other approaches have been used to reduce the metastability errors even further and to make the converter degrade gracefully in the presence of an error.

One possible approach consists in introducing pipelined latches immediately after the comparators outputs and before the logic decoder, increasing the regeneration gain of the comparator [Zoj 85], [Man 90], [Rey 94]. However, each pipeline stage needs $2^N - 1$ latches, increasing the die area and power consumption.

Another approach consists in using logic gates to convert the thermometer code directly to a binary code. By using the output of every comparator only once and by ensuring that the signal path never branches, an undecided comparator output only affects one output bit, which can be regenerated at the converter output [Por 96]. However, this approach implies that signals at intrinsically distant positions are logically associated. Typically, the wiring capacitance and the decoder area becomes unacceptable even in 6 bit designs.

The last approach consists in using ROM encoders combined with Gray codes. The situation illustrated in Figure 6.10(a) for the case of a binary coded ROM, is redrawn in Figure 6.10(b) for the case of a Gray coded ROM. In the case of a binary coded ROM, due to the two undecided selection lines, all bits from the binary coded ROM are undecided, and the final output word is in final analysis determined by random phenomena. Unfortunately, the same can happen with a Gray coded ROM. The two undecided selection lines yields three different possibilities: zero, one, or two ROM lines might be selected. If one or two lines are selected, the fact that the ROM is Gray coded solves the problem, since only one bit is undecided and it can be regenerated at the ROM output². However, if none of the ROM lines is selected, the ROM output bits are all one since they are determined by the ROM pull-up networks, leading to a severe error in the output code.

It can be concluded from the previous discussion that a Gray coded ROM is an effective solution to reduce the metastability errors caused by an undecided

²Since the input signal level is between the two comparator levels, it is not important if the undecided output bit is regenerated to a zero or a one.

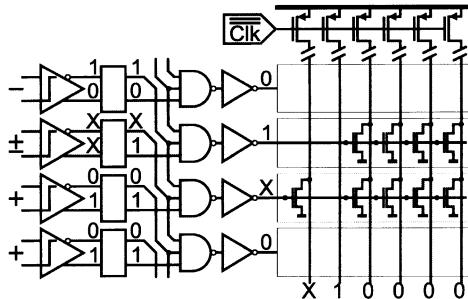


Figure 6.11: Avoiding metastability errors with a Gray coded ROM and a circuit that always ensures one correct logic level.

comparator output provided that the case of having no lines selected is eliminated by design.

Consider that the comparator is in a metastable situation—the two outputs are undecided. If these comparator outputs are applied to an auxiliary circuit that generates a valid output (a logic one) and an undecided output, and these two signals are then used to drive the decoder, then the problem is solved. This situation is graphically illustrated in Figure 6.11. The operation of this circuit is as follows. If the two outputs of the circuit are high when the ROM is clocked, two adjacent ROM lines are selected. Therefore, if a Gray coded ROM is used, the two words only differ in one bit, and the ROM output is the code with more zeros. If the two outputs of the circuit are complementary (like in a metastable free situation), only the correct ROM line is selected, and there is no problem. Finally, if the one of outputs of the circuit is settling to zero when the ROM is clocked (the other output is a valid one), then one ROM line is selected, and one ROM line can be in an invalid level. The selected line generates the correct output, while the other line can only alter the state of one output bit (assuming a Gray coded ROM), which can be regenerated at the ROM output.

In short, by ensuring that one of the signals driving the logic decoder is always at a valid one, a Gray coded ROM properly handles all the situations resulting from metastable comparator outputs.

Figure 6.12(a) shows the circuit presented in [Por 96] to convert the metastable input (the comparator outputs), in valid high logic levels. When the circuit inputs are equal, the top transistors are on and hold the circuit outputs at V_{DD} . When the difference between the input voltages is higher than the threshold voltage of the NMOS transistors, one of the outputs is pulled to low while the other remains at V_{DD} .

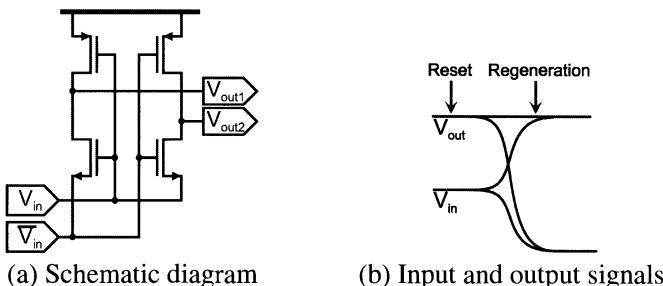


Figure 6.12: Possible implementation of the circuit necessary to avoid metastability errors.

In Figure 6.12(b) the described behavior is graphically illustrated by presenting the typical signals at the input and output of this circuit.

In this design two simple inverters have been used to accomplish the same effect—see Figure 6.4, transistors M_7 and M_8 . To understand why, consider the following argument. The ratio between the PMOS and the NMOS transistor used in these inverters is made higher than the ratio between the PMOS and NMOS transistors used in the regeneration loops. Consequently, the threshold voltage of the inverters (about 1.5 V) is intrinsically higher than the threshold voltage of the comparator (about 1.0 V). This means that when the comparator is reset, the two comparator outputs settle at the metastable point, and therefore the output of the inverters is a valid logic one. During the regeneration phase, if the comparator remains undecided, the two inverter outputs remain at the logic level one. When the comparator decides, one of the output changes to V_{SS} and the other to V_{DD} . Then, only the output changing to V_{DD} changes the corresponding inverter output to a logic zero.

Concerning the design of the remaining logic, the following details are worth mentioning.

The nand gate used in the logic decoder normally has a clocked NMOS and PMOS transistor that isolate the gate output from changes at its input during the regeneration phase. Since the two inverter outputs are always at the logic level one during the regeneration phase, the PMOS transistor has been omitted in this design.

For very high acquisition speeds, the voltage swing at the ROM output lines might be relatively small. Therefore, the ROM NMOS transistors have been drawn with a drain region surrounded by gate so that the total drain capacitance presented by the 32 transistors connected in parallel at each of the 6 bit lines is minimum. The

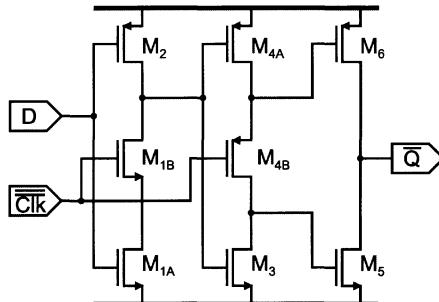


Figure 6.13: Dynamic latch used to buffer the ROM output codes during a complete clock period.

NMOS pull-down transistors used in the ROM have an W/L of about $4 \mu\text{m}/0.35 \mu\text{m}$, and the PMOS pull-up transistors have a W/L of $16 \mu\text{m}/0.35 \mu\text{m}$.

Figure 6.13 presents the dynamic latch used to converter the ROM output codes in digital signals valid during a complete clock period.

The next section presents the main experimental results obtained with the fabricated converter.

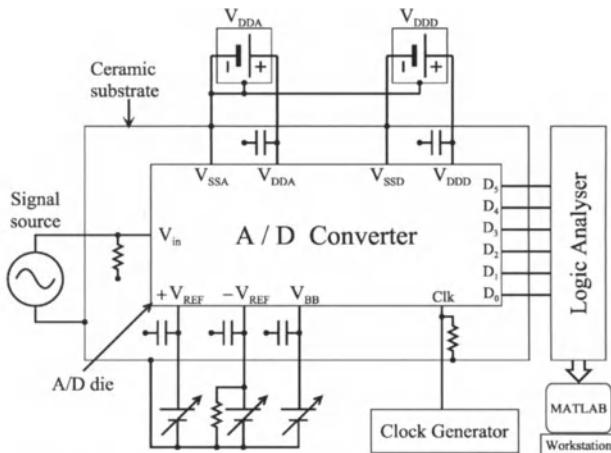
6.4 EXPERIMENTAL RESULTS

The converter presented so far has been implemented in a standard $0.35 \mu\text{m}$ CMOS technology process with double poly and four metal layers. The experimental results obtained with the fabricated converter are now presented.

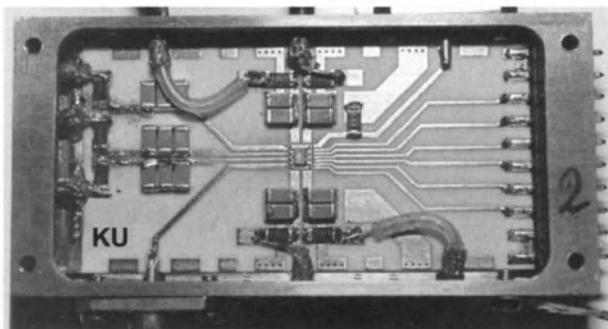
Figure 6.14(a) shows a schematic diagram of the measurement setup used in the performance evaluation of the flash A/D converter. To ensure an adequate decoupling of all DC lines, the converter has been mounted on a ceramic substrate where all supply and biasing lines are locally decoupled with SMD capacitors. A photograph of the measurement setup showing in detail the die mounted on the ceramic substrate and encapsulated on a shielding box is shown in Figure 6.14(b).

Figure 6.15(a) presents a microphotograph of the fabricated die, and Figure 6.15(b) shows the converter floor plan. From the bottom to the top one can see the ladder network, the preamplifiers, the comparators, the logic decoder, and the ROM. Each of the 63 sections of the flash converter are disposed next to each other. The converter active area is $0.9 \text{ mm} \times 0.6 \text{ mm}$, and the total die area (including bonding pads and decoupling capacitance) is $1.6 \text{ mm} \times 1.1 \text{ mm}$.

Figure 6.16(a) presents the spurious free dynamic range (SFDR) of the converter measured as function of the acquisition frequency for three different signal frequen-



(a) Schematic diagram of the experimental arrangement used for measurements.



(b) Photograph of the die mounted on a ceramic substrate inside of a shielding box.

Figure 6.14: Setup used in the measurement of the Flash A/D converter.

cies with a full scale amplitude. Figure 6.16(b) presents the total signal to noise plus distortion ratio (measured over the complete Nyquist band) for the same conditions as before.

The performance of this converter degrades gracefully with the sampling frequency and at about 950 MS/s it stops working³. The reason for the degradation when applying higher acquisition speeds is probably due to the limited comparators

³ According to simulation results the logic and gate does not function correctly above this frequency.

overdrive recovery strength. This clearly shows the need for good modelling of the comparator if one wants to take the 1 GHZ barrier (this will be done in the second implementation).

As can be seen from the SFDR plot, the performance of the converter also drops when the input signal frequency is increased for a fixed clock frequency. Several reasons can explain this: First, the coupling of the input signal to the ladder network is larger than expected, disturbing the reference voltages and so bad comparison are made. Second, the strong non linearity of the preamplifier input capacitance originates too much distortion. The total input capacitance of the flash ADC together with the signal source output impedance and termination resistance limits the analog input bandwidth. Even worse is the fact that the input capacitance is strongly voltage dependant—the number of preamplifier input transistors in saturation and in cut-off depends on the input voltage applied—, which originates harmonic distortion of the applied input signal. In other words, the input signal is subject to smaller delays for lower input voltages and to bigger delays for higher input voltages.

Maybe it is important to mention that measuring such high-speed analog designs is not trivial. One has to be able to apply a good and clean clock (with very low jitter) and also one has to generate high-frequency input signals with at least 6 bit accuracy otherwise not only the converter is measured but also the test equipment. Therefore a lot of filtering is needed to clean-up the clock and the input signals before applying them to the high-speed A/D Converter.

The current consumption in the ladder network is 70 mA. The total analog and digital current consumption are 230 mA and 50 mA, respectively (at an acquisition speed of 1000 MS/s). The converter has a total power consumption of 950 mW, and operates from a single power supply of 3.3 V. The main characteristics of the converter are summarized in Table 6.1.

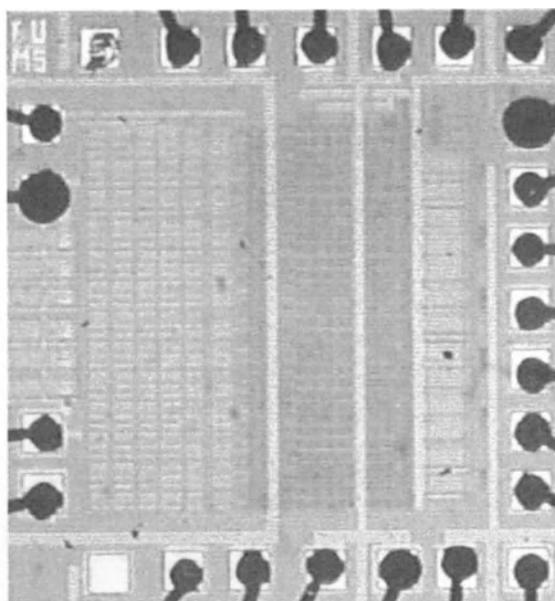
Process	0.35 μm CMOS
Input range	1.6 V
Resolution	6 bit
Maximum acquisition speed	1000 MS/s
SFDR (at 500 MS/s for a 241 MHz signal)	30 dB
SNDR (at 500 MS/s for a 141 kHz signal)	30 dB
SFDR (at 1000 MS/s for a 450 MHz signal)	25 dB
SNDR (at 1000 MS/s for a 141 kHz signal)	28 dB
Ladder resistance	100 Ω
Power consumption	950 mW
Chip area (converter core)	0.6 \times 0.6 mm ²
Chip area (with bondpads and decoupling)	1.22 \times 1.3 mm ²

Table 6.1: Main characteristics of the flash ADC.

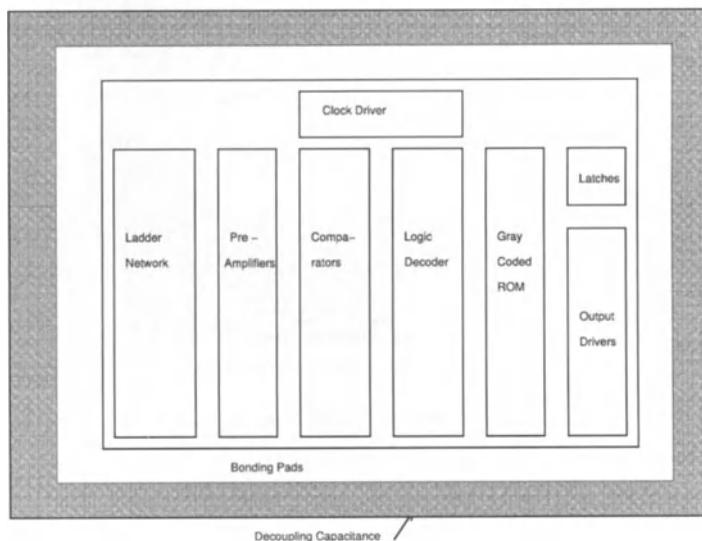
6.5 CONCLUSION

In this chapter, the design and measurement results of a 6 bit flash ADC has been presented. The circuit design of every building block has been discussed in detail. The implemented converter has been fabricated in a standard $0.35 \mu\text{m}$ CMOS technology. The converter core occupies an area of only 0.54 mm^2 . The maximum acquisition speed is 1000 MS/s. For low frequency input signals (and at the maximum acquisition speed), the SNDR is always higher than 31 dB. The total power consumption of the converter is 950 mW.

In the context of multi bit oversampled converters, the signal bandwidth and the required resolution of the multi bit quantizer are relatively small. On the other hand, the acquisition speed is normally to be maximized. The measurement results show that the former specifications, and a very high acquisition speed can be reached in standard CMOS technologies.

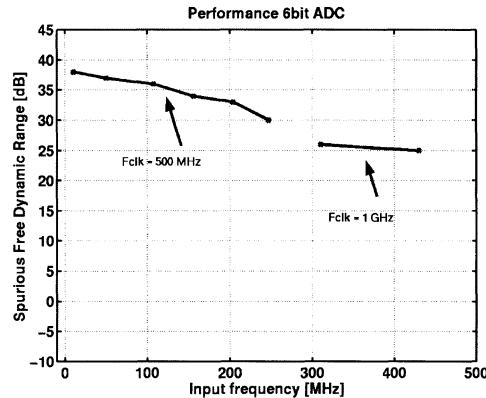


(a) Microphotograph of the flash A/D converter.

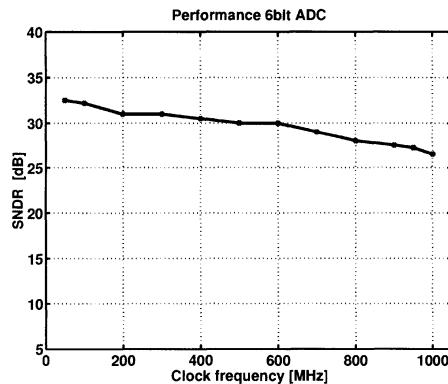


(b) Floor plan of the flash A/D converter.

Figure 6.15: Realized 6 bit Flash A/D converter.



(a) SFDR vs. acquisition speed.



(b) SNDR vs. acquisition speed

Figure 6.16: Converter performance as a function of the acquisition speed for different signal frequencies.

References

- [Alt 01] S. Altekar, “A 700-Mb/s BiCMOS Read Channel Integrated Circuit”, in *Proceedings International Solid-State Circuits Conference*, San Francisco, Feb. 2001, pp. 184–185.
- [Gar 89] V. Garuts, Y. Yu, E. Traya, and T. Yamaguchi, “A Dual 4-bit, 2-GS/s Full Nyquist ADC Using a 70-ps Silicon Bipolar Technology with Borosenic-Poly Process and Coupling-Base Implant”, *IEEE Journal of Solid-State Circuits*, vol. 24, no. 1, pp. 216–222, Apr. 1989.
- [Kae 97] F. Kaess, R. Kanan, B. Hochet, and M. Declercq, “New Encoding Scheme for high-speed Flash ADCs”, in *Proceedings IEEE International Symposium on Circuits and Systems*, Jun. 1997, pp. 878–882.
- [Man 90] C. W. Mangelsdorf, “A 400MHz Input Flash Converter with Error Correction”, *IEEE Journal of Solid-State Circuits*, vol. 25, no. 1, pp. 184–191, Feb. 1990.
- [Pel 89] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, “Matching Properties of MOS Transistors”, *IEEE Journal of Solid-State Circuits*, vol. SC-24, no. 3, pp. 1433–1439, Oct. 1989.
- [Pel 98] H. T. M. Pelgrom and M. Vertregt, “Transistor Matching in Analog CMOS Applications”, in *IEEE IEDM Techn. Dig.*, 1998, pp. 915–918.
- [Por 96] C. L. Portmann and T. H. Y. Meng, “Power-Efficient Metastability Error Reduction in CMOS flash A/D converters”, *IEEE Journal of Solid-State Circuits*, vol. 31, no. 8, pp. 1132–1140, Aug. 1996.
- [Rey 94] H. Reyhani and P. Quinlan, “A 5 V, 6-b, 80 MS/s BiCMOS flash ADC”, *IEEE Journal of Solid-State Circuits*, vol. 29, no. 8, pp. 873–878, Aug. 1994.
- [Uyt 00] K. Uyttendhoove and M. Steyaert, “A 6 Bit CMOS Very High Acquisition Speed Flash ADC with Digital Error Correction”, in *Proceedings Custom Integrated Circuits Conference*, Orlando, May 2000, pp. 432–436.
- [Ven 96] A. Venes and R. J. V. de Plassche, “An 80MHz 80mW 8b CMOS Folding A/D Converter with Distributed T/H Preprocessing”, in *Proceedings*

- International Solid-State Circuits Conference*, San Francisco, Feb. 1996,
pp. 241–243.
- [Zoj 85] B. Zojer, R. Petschacher, and W. A. Luschnig, “A 6-bit/200-MHz full Nyquist A/D converter”, *IEEE Journal of Solid-State Circuits*, vol. SC-20, no. 3, pp. 780–785, June 1985.

Chapter 7

Logarithmic Analogue-to-Digital Converters

Jorge Guilherme and Joao Vital¹

7.1. INTRODUCTION

Logarithmic signal processing is a topic that exists in the communications engineer's vocabulary for a long time. Sometimes referred as compression, it is a technique to adapt the dynamics of signals to the dynamics of the system where it is to be processed or transmitted. Probably, the most common application of compression is the telephone system, where the voice signal is compressed to a digital 8-bit signal [1-3]. Other applications such as critical military radar systems in electronic warfare, sonar/ultrasound and instrumentation applications requiring the compression of wide-dynamic-range signals, use common logarithmic amplifiers [4-7].

In recent years, special attention has been put in a particular class of compression applied to integrated filters, so called Log-Domain Filter Circuits [8,9]. In this case the main idea is also to compress the signal to be processed before noise and/or distortion have the chance to corrupt it. Due to compression (also known as companding), the voltage swing inside the log-domain filter is very small. The impedance level along the signal path is thus typically low, making the log-domain filters suitable for high frequency operation [10].

Logarithmic A/D converters also provide a non-uniform quantization, thus compressing the input analog signal to the digital world. They are widely used in communications, instrumentation and hearing aids, among other applications areas [11]. Originally, logarithmic converters have been implemented in bipolar technology to explore the inherent exponential I-V characteristic of the transistors [12]. Alternative architectures for implementation in the more cost-effective CMOS technology have also been proposed based on the piecewise approximation of a logarithmic function [13-15]. There is a wide range of

¹ Jorge Guilherme is with Escola Superior Tecnologia Tomar, and João Vital is with Chipidea – Microelectrónica, S.A. and with Instituto Superior Técnico, Universidade Técnica de Lisboa, Portugal

different topologies to realize non-uniform quantization. The description of all the possibilities is beyond the scope of this chapter and will not be referred here.

This chapter describes a new type of logarithmic pipeline A/D converter suitable for high-frequency applications, where the signal operations are carried out in the logarithmic domain by simple scaling operations. A new symmetrical logarithmic A/D converter employing a 1.5bit/stage pipeline architecture suitable for implementation in CMOS technology is described. Digital error correction and calibration techniques traditionally used in linear converters are extended to the logarithmic case.

7.2. NON-UNIFORM QUANTIZATION METHODS

Non-uniform quantization employs code transition levels that are not equally spaced over the full input range, as illustrated in Figure 7.1. High-level input values are coarsely quantized and low input level values are quantized with the maximum accuracy. The realization of a non-uniform quantizer can be implemented by means of pre-compression, and post-expansion of the signal, and the architecture is called a companding converter [12]. The other way of implementing non-uniform quantization levels is obtained by comparing the input signal with non-uniformly spaced reference levels [16]. The set of quantization levels can follow different conversion laws, as exponential [17] or decimal factors [18].

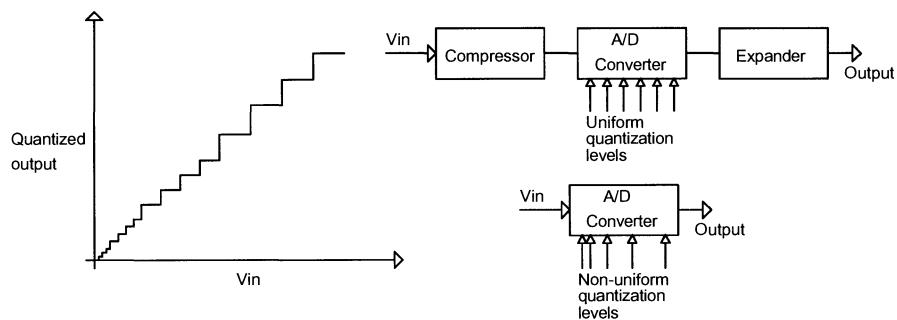


Figure 7.1: Non-uniform quantization.

One of the most common laws used is the μ -law, shown in Figure 7.2 [2]. Its conversion characteristic is given by:

$$y = \text{sign}(V_{in}) \frac{\ln(1 + \mu |V_{in}|)}{\ln(1 + \mu)} \quad (7.1)$$

where y is the digital normalized output, μ the compression coefficient and $\text{sign}(x)$ is the sign function.

The logarithmic conversion in this case, is done by first converting the input signal with a linear ADC with 13-bits of resolution. Then, the output of the ADC is used to address a lookup-table from which an 8-bit compressed code is fetched. Table 7.1 shows the input-output characteristic, considering an input digital 13-bit code and an 8-bit output digital code.

Table 7.1: μ -law input-output characteristic.

Linear code	Compressed code
00000001wxyzab	000wxyz
00000001wxyzabc	001wxyz
000001wxyzabcd	010wxyz
00001wxyzabcde	011wxyz
0001wxyzabcde	100wxyz
001wxyzabcdef	101wxyz
01wxyzabcdefg	110wxyz
1wxyzabcdefg	111wxyz

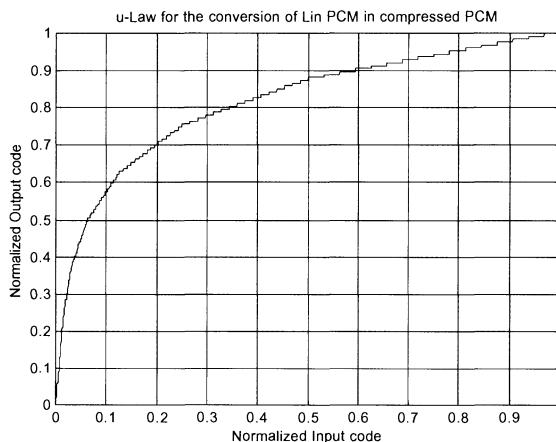


Figure 7.2: Half characteristic of a standard CCITT G.711 μ -law converter.

7.2.1 Low speed non-uniform quantizers

In the low-speed range of non-uniform quantizers there are various kinds of converter architectures, the most common using the pulse-width modulation (PWL) scheme, the sigma-delta approach, the ramp type topology, or the successive approximation algorithm, among others [19-24]. The description of all of the different types of possible topologies is beyond the scope of this chapter. Therefore only two of the most used topologies will be described.

The first to be addressed is the ramp type, illustrated in Figure 7.3. In this converter the input voltage is coded according to the transfer function stored in a ROM (or EPROM), which can be arbitrarily chosen [19]. The conversion is performed by feeding the ROM contents sequentially into the DAC until its output equals the input voltage. During that time the counter counts the number of clock pulses. The total number of pulses stored in the counter is the representation of the input voltage according to the ROM conversion law.

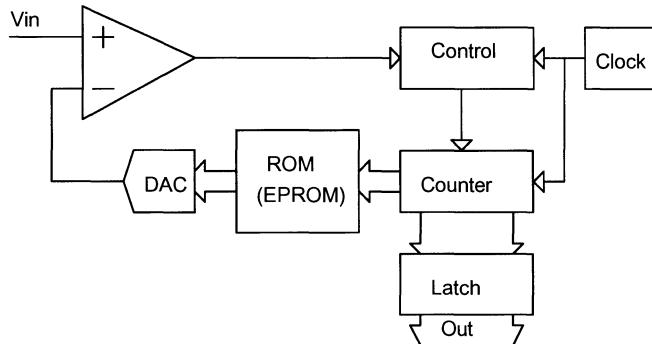


Figure 7.3: Ramp type non-linear A/D converter.

The successive approximation type ADC is illustrated in Figure 7.4. The principle here is to find an approximation of the input voltage by successive attenuations of a reference voltage with a programmable attenuator. The conversion algorithm is controlled by the use of a successive approximation register (SAR) and a comparator [13-15]. When the output of the attenuator reaches the closest value to the input voltage, the digital control word of the attenuator is proportional to the logarithm of the unknown input voltage.

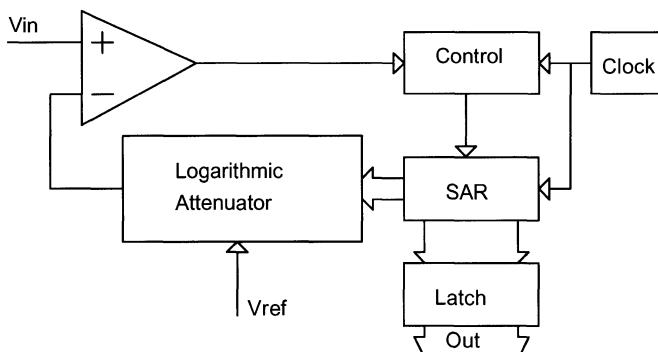


Figure 7.4: Block diagram of a successive approximation logarithmic converter.

A partial realization of this method is found in autoranging digital voltmeters, which determine the decade in which the unknown voltage lies by finding out

how many amplifications or attenuations by a factor of ten are required to bring it into a standard decade [15].

7.2.2 High-speed non-uniform quantizers

High-speed operation is achieved by the use of the flash and pipeline architectures [16, 24, 26]. The most straightforward way of implementing a non-uniform quantizer is to use the flash architecture where the resistor string voltage taps have a non-uniform spacing, as shown in Figure 7.5 [26]. The resistor string values can be calculated by [16]

$$R_n = R \left[V_T(n) - V_T(n-1) \right] \quad \text{with} \quad R = \sum_{n=1}^N R_n , \quad n = 1, \dots, N \quad (7.2)$$

$$R_n = R \cdot V_T(N) \quad (7.3)$$

and $V_T(n)$ are the non-uniform comparator threshold voltages.

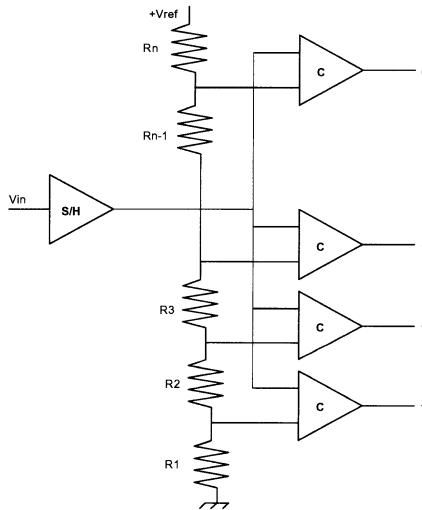


Figure 7.5: Flash architecture with non-uniform quantization.

The threshold voltages can be designed for any given conversion law. In this way, the output of the comparators is a compressed representation of the input signal according to the coded conversion law. This principle can be applied to other topologies such as the two-step flash [24] and the pipeline, which will be the subject of the next sections.

Recently a special type of non-uniform quantizer, so-called floating-point ADC (FADC) [27], is receiving special attention. These converters maintain a

linear relationship between the input voltage and the digital output by the use of floating-point arithmetic, which is commonly used in digital signal processors (DSP). The floating-point representation is found using a variable gain amplifier, or selectable gain amplifier, that controls the input amplitude voltage of a normal linear A/D converter [27-30], as illustrated in the block diagram in Figure 7.6. Other possible architectures exist, but will not be discussed here.

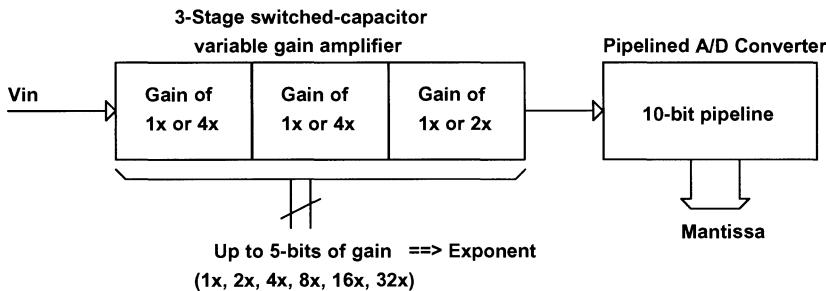


Figure 7.6: Block diagram of a floating-point A/D converter [27].

The output of the converter is given in a binary floating-point number as

$$Data = sign(Vin) \cdot m \cdot 2^c \quad (7.4)$$

with m being the mantissa and c the exponent.

The exponent is the digital word that controls the gain of the variable amplifiers. The mantissa is the output of the pipelined A/D converter. To get the correct output value a simple bit-shift operation of the mantissa according to the exponent factor is needed to reconstruct the signal.

This converter as the other cases mentioned in this chapter maintains a constant relative error in the output digital data, rather than an absolute error as in traditional linear converters. The dynamic range is independent of the maximum signal-to-noise ratio (SNR), which is dependent on the number of bits used in the mantissa pipeline ADC. The number of bits in the exponent determines the maximum dynamic range of the converter.

7.3 TRANSFER CHARACTERISTIC

The transfer characteristic of a logarithmic A/D converter can be defined by the following dual relationship (7.5), normalized to 1 V input, for positive values only. In this expression K is the compression coefficient, N the number of bits and D_n the digital output between 0 and 2^N-1 . Figure 7.7 shows the conversion characteristics for some compression coefficient values [31-33].

$$\begin{cases} V_{in} = K \cdot \left(1 + \frac{1}{K}\right)^{\frac{D_n}{2^N - 1}} \\ D_n = (2^N - 1) \cdot \frac{\ln\left(\frac{|V_{in}|}{K}\right)}{\ln\left(1 + \frac{1}{K}\right)} \end{cases} \quad (7.5)$$

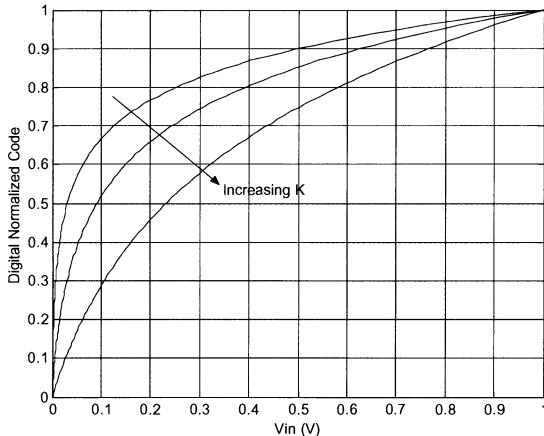


Figure 7.7: Normalized conversion characteristic of a logarithmic ADC.

When the compression factor K decreases, the dynamic range increases, i.e. the ratio of the maximum input signal range to the minimum transition step increases. The logarithmic converter can be designed to have the dynamic range independent of the number of bits, since the compression factor K and N are independent from each other. The minimum threshold voltage in (7.5) defines the LSB of the converter, as given by (7.6). The dynamic range is calculated by the ratio of the input range to the LSB of the converter and is given in (7.7), considering only positive input signals.

$$V_{LSB} = K \cdot \left(\left(\frac{1}{K} + 1 \right)^{\frac{1}{2^N - 1}} - 1 \right) \quad (7.6)$$

$$DR_{dB} = 20 \cdot \log_{10} \frac{1}{K \cdot \left(\left(\frac{1}{K} + 1 \right)^{\frac{1}{2^N - 1}} - 1 \right)} \quad (7.7)$$

7.4. SIGNAL-TO-NOISE RATIO

The maximum signal-to-noise ratio (SNR) at the output of a logarithmic A/D converter for a sinusoidal input is given by [1, 34]

$$\text{SNR}_{\text{dB}} = 10 \cdot \log_{10} \frac{3 \cdot 2^{2N}}{\left[\ln \left(1 + \frac{1}{K} \right) \right]^2} \quad (7.8)$$

The number of bits and the compression factor K can be used to achieve the required resolution and dynamic range.

An ideal N -bit linear converter has a peak SNR given by (7.9), considering only the quantization noise, and assuming that the input signal is a sinusoid with full-scale amplitude.

$$\text{SNR}_{\text{dB}} < 6.02N + 1.76 \text{ dB} \quad (7.9)$$

In such cases where the input signal does not reach full-scale, the resulting SNR is given by

$$\text{SNR}_{\text{dB}} = 1.76 + 6.02N + 20 \log_{10} \left(\frac{V_{\text{amplitude}}}{V_{\text{MAX}}} \right) \quad (7.10)$$

where $V_{\text{amplitude}}$ is the peak-to-peak signal amplitude and V_{MAX} is the converter input range. For gaussian random input signals the peak SNR is 9 dB below the value given by (7.9) [41,42].

Figure 7.8 shows the statistical computer simulated SNR of a symmetrical 9-bit, 80 dB ($K = 0.01$) dynamic range pipeline A/D converter together with its linear counterpart, for which different resolutions between 11 and 13 bit were considered. The simulations were performed with white noise input signals with more than 10.000 points for each input voltage amplitude step. From this Figure we can conclude that a 9-bit logarithmic converter is superior to a 12 bit linear ADC in the lower range of input signal amplitude. The maximum simulated SNR for the logarithmic ADC is 44.3 dB. A slightly better result of 45.67 dB would be obtained for sinusoidal input signals, as predicted by equation (7.8).

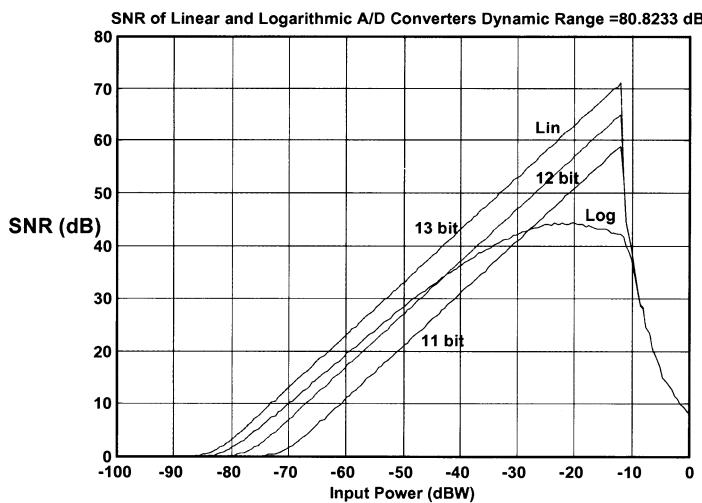


Figure 7.8: Statistical computer simulation of the *SNR* of a 9 bit logarithmic converter, compared to the SNR for 11, 12 and 13 bit linear converters.

One possible application of the logarithmic converters is in data communications requiring a wide dynamic range to accommodate the dynamics of the received signal, but that do not necessarily require a high SNR. An example of such communications systems is ADSL. The modulation method employed, 14 bit QAM, only needs about 42 dB of SNR to achieve a Bit Error Rate (*ber*) of 10^{-8} , as shown in Figure 7.9.

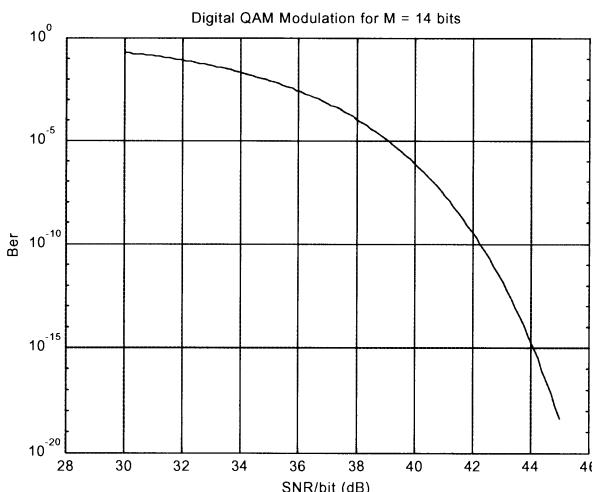


Figure 7.9: Probability of bit error as a function of SNR for 14 bit QAM.

This suggest that a converter with a nonlinear characteristic achieving an SNR of 42 dB and a dynamic range of 80 dB can be used in this type of applications, with the advantage of the reduced number of bits. The logarithmic converter has also the advantage that the output SNR is independent of the statistics of the input signal [1, 41, 42]. This is ideal for the case of a DMT signal used in ADSL, which consists of a sum of M independent quadrature modulation (QAM) signals, referred to as subchannels [1]. For large values of M ($M > 10$), the ensemble average of the instantaneous amplitude $A(t)$ can be accurately modeled by a Gaussian random process (central limit theorem) with a zero mean and a variance σ^2 equal to the total signal power [1, 35]. This type of signal is well modeled by a white noise source, similar to what was used to obtain the results presented in Figure 7.8.

7.5. PIPELINE LOGARITHMIC ADC

The next two sections will present a design methodology that can be applied to any converter topology. This methodology consists on a mathematical transformation of the signal operations that are performed in the linear domain on conventional ADCs to those in a logarithmic domain. The result of this transformation modifies almost any linear converter in a logarithmic one. The most interesting topologies to apply this transformation are the pipeline and the two-step flash, due to their high-speed operation. The flash topology is excluded from these considerations since in this topology the input signal is not processed, but rather only compared to a given set of threshold voltages. The design methodology will be illustrated by two examples: a 1.5-bit/stage pipeline ADC and a two-step ADC.

7.5.1 Logarithmic transformation

The conceptual block diagram of a classical, linear pipeline ADC converter with 1.5b/stage is shown in Figure 7.10. Each stage of the pipeline converter operates in a straightforward way. During the first quantization cycle the flash converter quantizes a sampled-and-held input into 1.5 digital bits. During the subsequent second quantization cycle, a reconstructed voltage corresponding to the converted bits is subtracted from the sampled-and-held input to generate a residue voltage. This residue voltage is amplified by 2 and then applied to the next pipeline stage. The converted 1.5 bits are passed through delay registers, to align all the bits corresponding to the same input sample, and then applied to the digital correction logic to extract the final output bits.

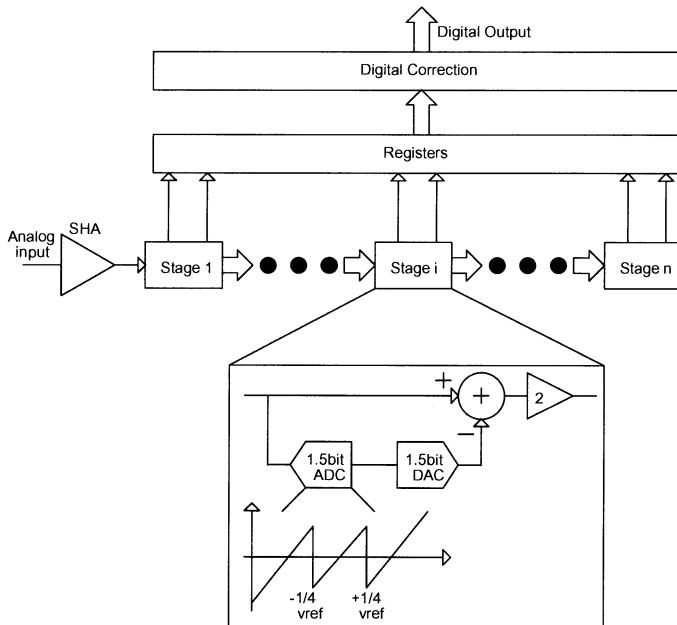


Figure 7.10: Block diagram of a 1.5b/stage pipeline A/D converter.

For the realization of a logarithmic pipeline A/D converter, the operations above have to be carried out in the logarithm domain, where the subtraction is equivalent to a division, and the multiplication by 2 is equivalent to a squaring operation. Such operations can be explained by representing the logarithmic transformation as a pair of functions, $f(x)$ and $g(y)$, were $g(y)$ is the inverse function of $f(x)$ as represented in Figure 7.11.

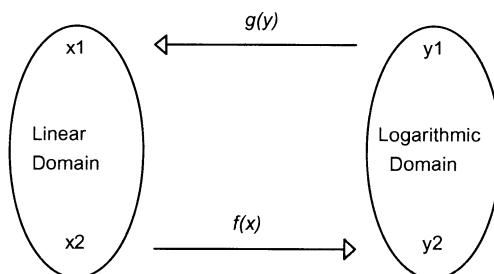


Figure 7.11: Relationship functions between the linear domain and the logarithmic domain.

The $f(x)$ and $g(y)$ functions are derived from (7.5) and can be represented by

$$\begin{cases} f(x) = K \cdot \left(\frac{1}{K} + 1 \right)^x = a \cdot b^x \\ g(y) = f(x)^{-1} = \frac{\ln\left(\frac{|y|}{K}\right)}{\ln\left(\frac{1}{K} + 1\right)} = \frac{\ln\left(\frac{|y|}{a}\right)}{\ln(b)} \quad \text{for } x = \frac{D_n}{2^N - 1} \end{cases} \quad (7.11)$$

where K is the compression coefficient, $a = K$ and $b = 1+1/K$. For $N = 8$ and $K = 0.01$ we obtain the conversion characteristic of Figure 7.12 [24]. The inherent logarithmic characteristic of these converters allows only positive input signals due to the *log* function. However positive and negative input signals can also be possible, as it will become clear in the next section.

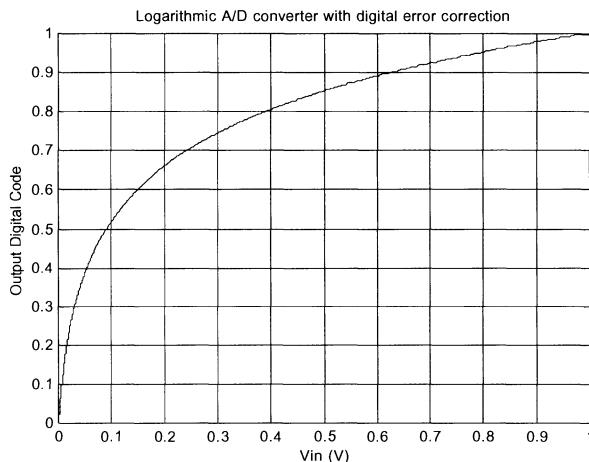


Figure 7.12: Example of a conversion characteristic of an 8-bit logarithmic ADC.

The flash converter must have a logarithmic quantization law to extract the 1.5 bits that control the DAC. Then, in the linear case the output of the DAC is subtracted from or added to the input, whereas in the logarithmic domain, where the subtraction is equivalent to a division and the addition to a multiplication, this operation corresponds to an attenuation or amplification that depends on the input digital word from the flash ADC. Thus the required signal operations after the subtraction are $x \pm \frac{1}{4}$ V, for a normalized positive input range of 1 V.

Based on this approach the DAC is replaced by a digitally controlled logarithmic attenuator whose attenuation factors are obtained by applying to $f(x)$ in (7.11) the correspondent output voltages of the subtractor. The signal transformation is $f(x \pm \frac{1}{4})$ that corresponds to an output of $f(x) \cdot b^{\pm\frac{1}{4}}$. Thus the coefficients are

$$A_0 = b^{\frac{1}{4}} \quad A_1 = 1 \quad A_2 = b^{-\frac{1}{4}} \quad (7.12)$$

A_1 corresponds to the gain in the middle segment of the 1.5-bit residue characteristic. It can be noted that the A_0 factor is the inverse of the A_2 factor, and this can be easily implemented by exchanging the input resistors or capacitors with those in the feedback paths in an inverting feedback amplifier configuration. The reference voltages of the flash converter are given by (7.13) after applying the threshold voltages of a 1.5-bit flash to $f(x)$ in (7.11), with only positive values ($x = 3/8$ and $x = 5/8$).

$$V_{ref1} = a \cdot b^{\frac{3}{8}} \quad V_{ref2} = a \cdot b^{\frac{5}{8}} \quad (7.13)$$

The reference voltage V_{ref1} can be obtained by multiplying A_2 with V_{ref2} .

The last stage of the pipeline is a 2 bit flash converter. Its reference voltages are obtained in the same way as before and are given by

$$V_{ref3} = a \cdot b^{\frac{1}{4}} \quad V_{ref4} = a \cdot b^{\frac{2}{4}} \quad V_{ref5} = a \cdot b^{\frac{3}{4}} \quad (7.14)$$

The multiplication by two in the linear case is replaced by a squaring operation and a scaling factor in the logarithmic domain. The output of the logarithmic attenuator is $f(x)$ multiplied by one of the coefficients from (7.12). The resulting voltage is squared, corresponding to residue amplification in the linear case. The scaling factor is needed since this squared voltage must correspond exactly to the amplification by two in the logarithmic domain. The scaling factor can be determined by noting that from $f(x)$ one LSB at 1.5-bit level is $ab^{\frac{1}{4}}$ and, after multiplication by A_0 and applying the squaring operation, the resulting voltage should be equal to half scale point, $ab^{\frac{1}{2}}$

$$\left[f\left(x = \frac{1}{4}\right) A_0 \right]^2 G_s = f\left(x = \frac{1}{2}\right) \quad (7.15)$$

Since this does not happen, it is necessary to introduce the scaling factor given by

$$G_s = \frac{1}{a \cdot \sqrt{b}} \quad (7.16)$$

7.5.2 Logarithmic pipeline ADC

The result of the above operations is a logarithmic pipeline converter with 1.5 bits/stage whose architecture is shown in Figure 7.13. An offset of a needs to be added to the input signal to allow the input voltage to go to zero. Figure 7.14 shows the residue voltage of one pipeline stage, where it is possible to observe the parabolic transfer curve of the squarer. All the stages have an identical topology, except the last one that resolves 2-bits with the flash ADC and does not need neither the squarer nor the attenuator.

This architecture can only process positive input signals, due to the logarithmic characteristic. The digital correction logic is exactly the same as the one used for linear converters [43].

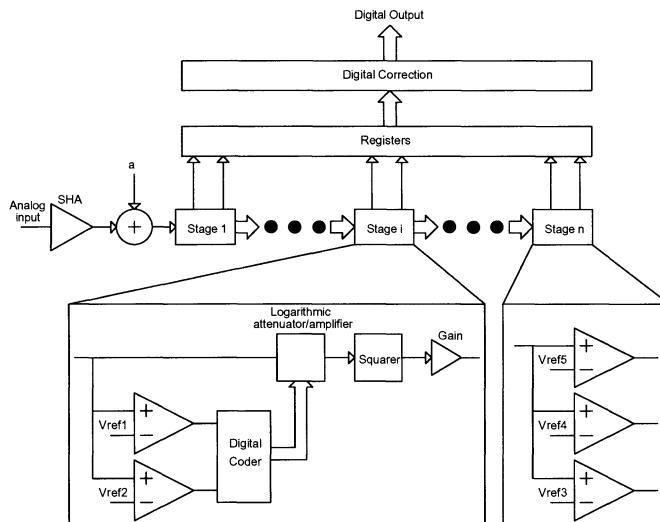


Figure 7.13: Block diagram of a logarithmic pipeline ADC with 1.5b/stage.

The flash comparators need to have an offset voltage of less than $a \cdot b^{\frac{1}{8}}$ (half LSB in the linear case). However by noting that the set of transition voltages on the flash ADC is not linearly spaced, the comparator connected to the higher threshold voltage can have also a higher offset and consume less power. For this

comparator the offset voltage needs to be lower than $a \cdot b^{\frac{1}{2}}(b^{\frac{1}{8}} - 1)$. That corresponds to the difference between V_{ref2} and half LSB ($a \cdot b^{\frac{5}{8}} - a \cdot b^{\frac{4}{8}}$).

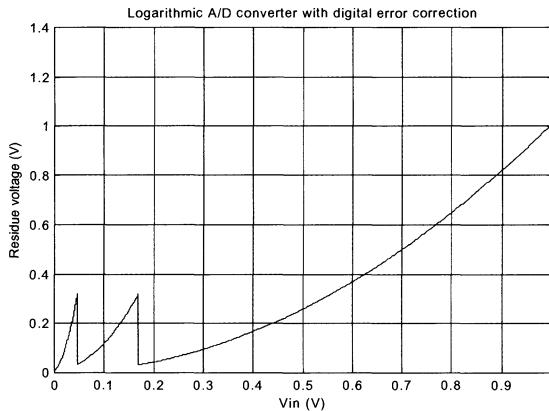


Figure 7.14: Residue voltage of one pipeline stage, with its parabolic transfer characteristic.

7.5.3 Symmetrical pipeline logarithmic ADC

The architecture shown in Figure 7.13 can only process positive input signals. To allow the use of bipolar signals a transformation in the input stage has to be done. Noting that after the first stage, the squarer rectifies all the positive or negative signals, the following stages will only process positive signals. The first stage will extract the most significant bits and the sign bit. Figure 7.15 shows the symmetrical characteristic of such an ADC. Figure 7.16 shows the resulting architecture for the symmetrical pipeline ADC, and Figure 7.17 shows the residue voltage at the output of the first stage.

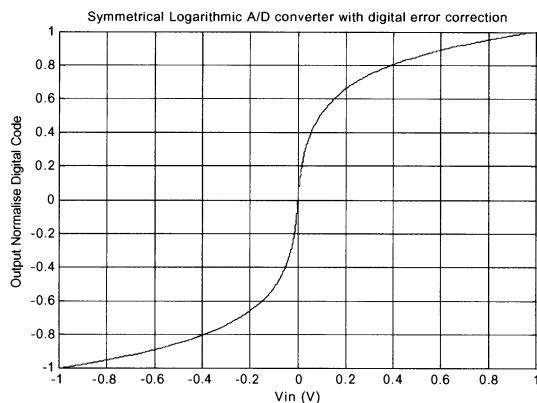


Figure 7.15: Symmetrical logarithmic converter characteristic.

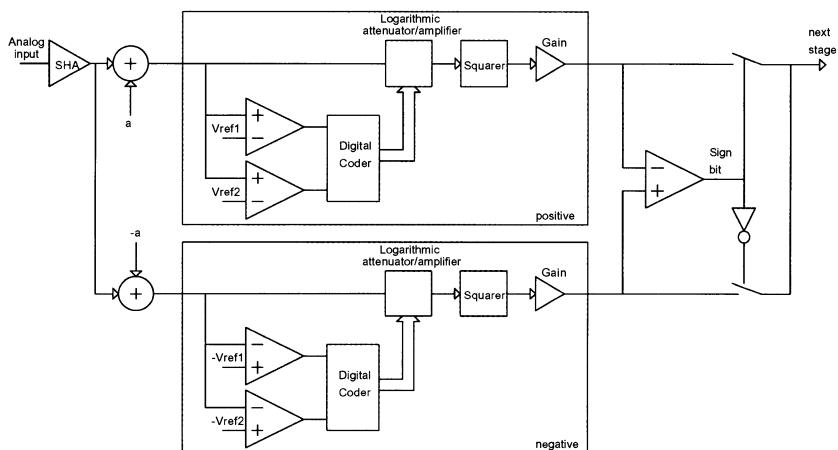


Figure 7.16: Input block diagram of the symmetrical logarithmic pipeline ADC with 1.5b/stage.

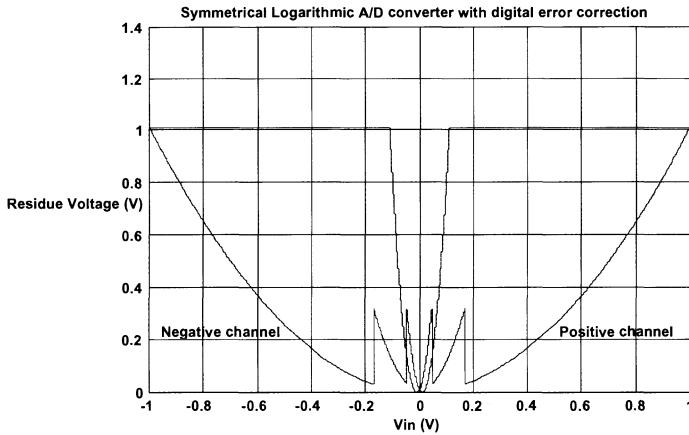


Figure 7.17: Residue voltage of first stage of the pipeline.

The sign bit is extracted comparing the output of both input stages. The sign comparator will detect the crossing of both output voltages. This method allows the use of a less precise comparator, since it will benefit from the voltage gain of the input stages. If this were not done, the comparator would need to have a resolution equal to the LSB of the converter, given by (7.6), [32, 33]. The comparator will need to have an offset voltage lower than half LSB of the second stage affected by the square root function. Since it will compare the outputs of the two channels, the difference voltage at its input will be doubled. The offset voltage is thus given by

$$V_{os} \leq K \left[\left(1 + \frac{1}{K} \right)^{\frac{1}{2^{N-1}-1}} - 1 \right] \quad (7.17)$$

7.5.4 Architecture improvements

The attenuation of the input signal is bad in terms of the noise performance of the converter, because it corresponds to an increase of the noise contribution from the last stages. The programmable attenuator has a gain smaller than one and, at the same time, the squarer has bad noise performance. The way to improve the noise performance of the system is to transfer the gain block after the squarer to the input. Figure 7.18 shows the modification in the pipeline stage.

The attenuation factors A_0 , A_1 and A_2 will be scaled by the square of the gain block.

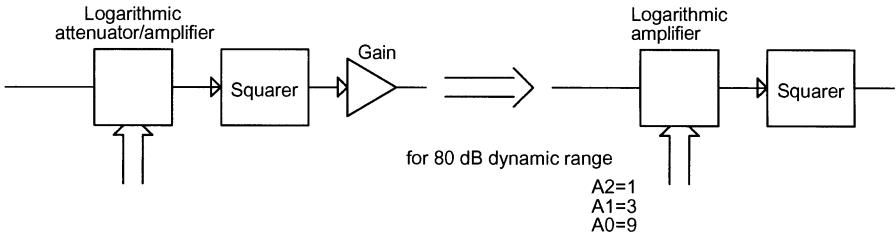


Figure 7.18: Substitution of the gain block to improve noise.

The new scaling factors with the gain block transferred to the input are

$$A_0 = a^{-\frac{1}{2}} \quad A_1 = \frac{1}{a^{\frac{1}{2}} \cdot b^{\frac{1}{4}}} \quad A_2 = \frac{1}{a^{\frac{1}{2}} \cdot b^{\frac{1}{2}}} \quad (7.18)$$

however, if one of this gains could be set to unity, this would simplify the circuit implementation. The A_2 gain is very close to unity, which suggests that this small gain difference can be included in the squarer and set A_2 to unity. Doing this alteration the squarer gain can be changed from unity to $1+a$. Normally $a \ll 1$, since a is equal to the compression factor ($a = K$). The modified scaling factors are then given by

$$A_0 = b^{\frac{1}{2}} \quad A_1 = b^{\frac{1}{4}} \quad A_2 = 1 \quad (7.19)$$

This alteration changed the programmable attenuator into a programmable gain amplifier (*PGA*), with a minimum gain of one. The small amplitude signals will be amplified by A_0 , thus reducing the input referred noise contribution of the squarer and following stages.

The symmetrical characteristic presents a new problem to extract the sign bit. The sign bit is extracted comparing the output of both input stages during the residue amplification phase. The sign bit will be ready only at the end of the residue amplification phase of stage one, which coincides with the sampling phase of stage two. This lead to the duplication of the *PGA* block in stage two. The sign bit is then used to select which *PGA* output is used as input of the squarer. This alteration will need one more *PGA* and 1.5-bit flash, increasing the power consumption. Figure 7.19 shows the final pipeline architecture with the proposed changes [33].

The signal amplitude is normalized to ± 1 V, therefore to extract the sign bit, it is better to compare the output of both PGAs since the signal amplitude will be higher than after the squarer (the square of a number less than unity is even lower than unity). The offset of the sign comparator must be less than

$$V_{os} \leq \sqrt{K} \left[\sqrt{\left(1 + \frac{1}{K} \right)^{\frac{1}{2^{N-1}-1}} - 1} \right] \quad (7.20)$$

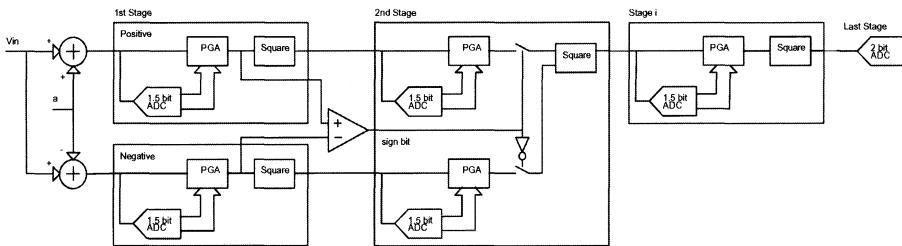


Figure 7.19: Final improved pipeline architecture.

7.6 LINEARITY CHARACTERIZATION

Linear A/D converters are usually characterized in terms of integral nonlinearity (INL) and differential nonlinearity (DNL) characteristics. The difficulty in a logarithmic converter is the dependence of the LSB on its position on the transfer function, not being constant. The LSB is constant only in the logarithmic domain.

Specifications in terms of INL and DNL are very rare and diffuse in the literature. The test standards used for ADCs do not cover the logarithmic or non-linear case [36]. Therefore, there is the need to correctly define INL and DNL with a coherent method such that these parameters can give a good representation of the converter characteristics.

If the same type of signal transformations used for the logarithmic converters are applied here, it is possible to define, in the logarithmic domain, the INL and DNL characteristics [37, 38]. By applying logarithms to $f(x)$, we get

$$\ln(f(x)) = \ln(a) + x \ln(b) \quad (7.21)$$

the nonlinear function $f(x)$ has been transformed in a linear function in the logarithmic domain. The transition step is given by the x coefficient, $\ln(b)$. Therefore the INL and DNL can be calculated by

$$\text{DNL} = \frac{\ln(V_{k+1}) - \ln(V_k) - \Delta V_{in}}{\Delta V_{in}} \quad (7.22)$$

$$\text{INL} = \frac{\ln(V_{k+1}) - \ln(V_i)}{\Delta V_{in}} \quad (7.23)$$

$$\Delta V_{in} = \ln\left(1 + \frac{1}{K}\right) \quad (7.24)$$

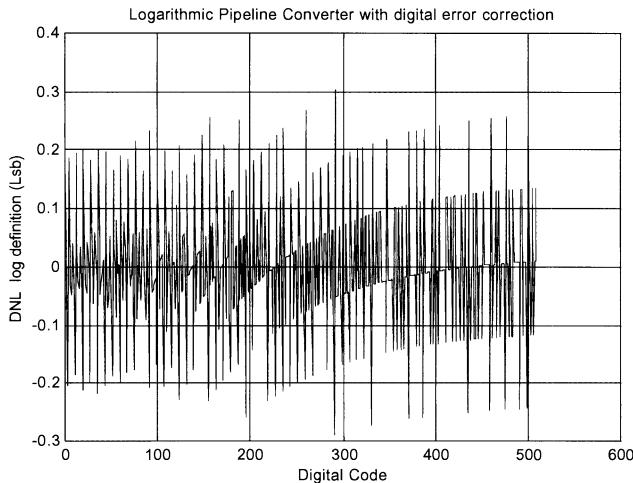


Figure 7.20: DNL in the logarithmic domain of the pipeline converter.

Figure 7.20 and Figure 7.21 show the DNL and INL characteristics for a MATLAB model simulation of a logarithmic pipeline ADC, with V_{in} varying from -1 V to 1 V, assuming a precision of 0.2% for the PGA coefficients, 15 mV offset for the comparators. This two Figures (7.20 and 7.21) shows the characteristics of the symmetrical pipeline logarithmic converter with $8 +$ sign bit resolution and 80 dB dynamic range.

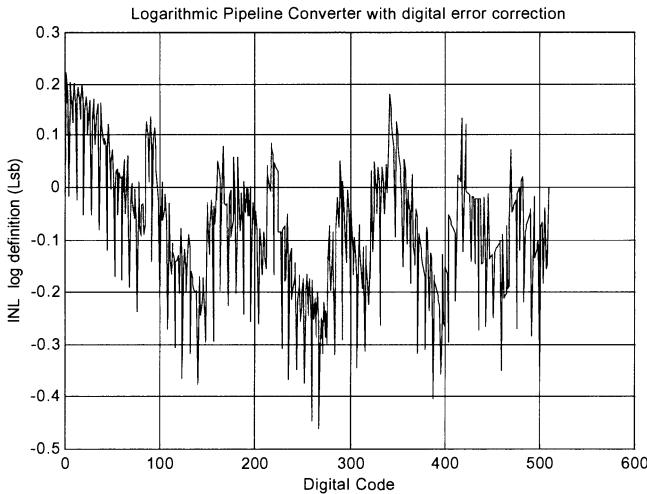


Figure 7.21: INL in the logarithmic domain of the pipeline converter.

7.7 DIGITAL CALIBRATION

Digital calibration similar to the one used in linear pipeline ADCs can be applied to the logarithmic converter. This calibration allows the cancellation of gain errors in the PGA and squarer blocks [39]. The developed digital calibration can compensate these gain errors as they are equivalent to an offset in the logarithmic scale. The calibration operates by aligning all the segments of the global transfer characteristic of the ADC.

The calibration procedure is started from the last but one stage and moves towards the first frontend stage. The pipeline chain formed by all the stages to the end of the pipeline are used as a quantizer to determine error codes. The generation of the error codes for each stage is performed by breaking the pipeline and by sequentially applying three input reference voltages to the input of the stage under calibration. The corresponding error codes are obtained by making a difference between the digital outputs of the remaining pipeline chain and the ideal ones. These errors are used to correct that stage and then the next one is calibrated having this one as a part of the remaining pipeline chain. The codes are stored in a RAM and will be accessed during the normal converter operation. The calibration input voltages are V_{ref3} , V_{ref4} and V_{ref5} , the reference voltages of the last stage.

During normal conversion the calibration code is fetched from RAM from the address pointed by the digital code that has been quantized by that stage, and added to the output. Figure 7.22 illustrates how the output word is obtained. Note

that an extra bit is considered to avoid problems due to digital truncation in the computations. This extra bit is dropped in the output word. Figure 7.23 and Figure 7.24 shows the INL and DNL MATLAB simulations of the converter with and without digital calibration for the extreme case of 5% precision in the *PGA* gains and 20% precision in the gain that affects the squarer ($G_S \times x^2$).

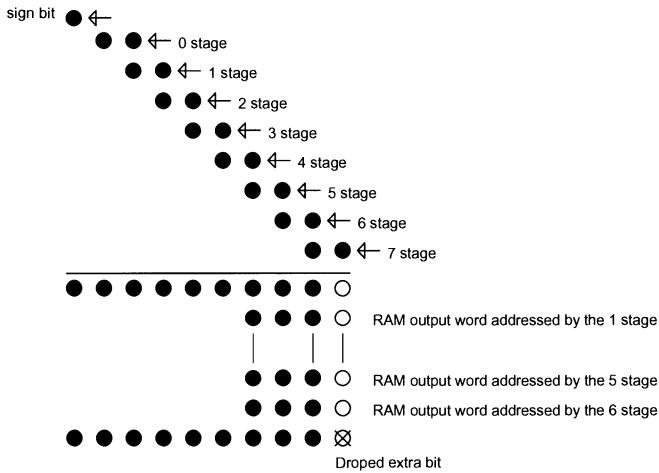


Figure 7.22: Combination of the intermediate quantization outputs and calibration codes from the output word.

The input stage has two channels, one for positive signals and the other for negative. Due to this parallelism, there will always be small gain differences between both channels. This gain mismatch can cause unwanted frequency components in the output. The digital calibration can correct the normal gain mismatch between both channels in the input stages. The sign bit is used by the algorithm to distinguish between positive or negative inputs and in that way it gives the possibility to calibrate each of the two input channels

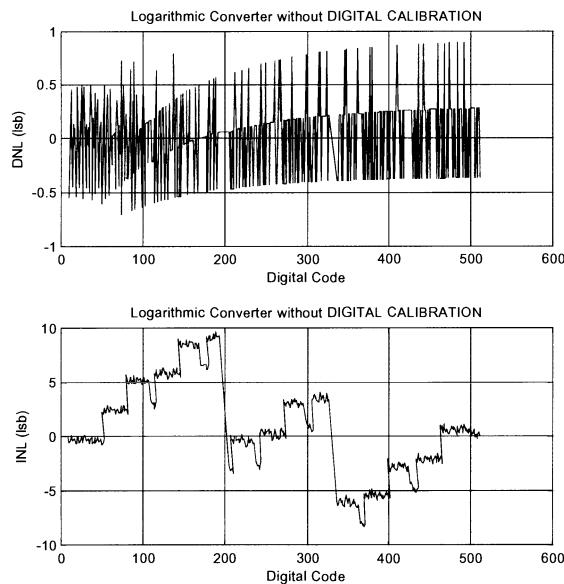


Figure 7.23: Matlab simulations without digital calibration for the case of 5% error in the attenuator gain and 20% error in the squarer gain.

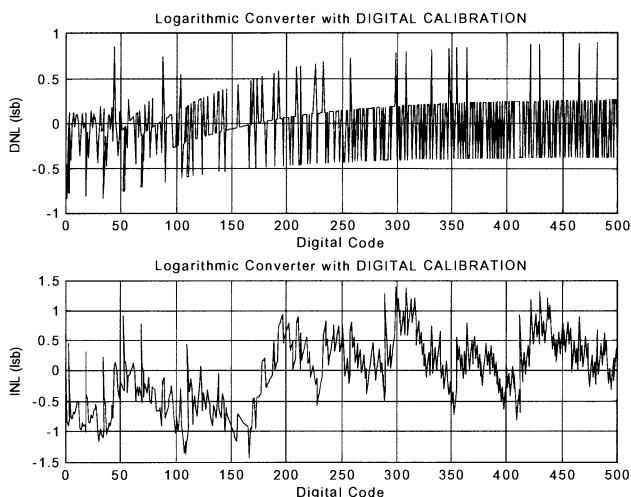


Figure 7.24: Matlab simulations with digital calibration for the case of 5% error in the attenuator gain and 20% error in the squarer gain.

7.8 TWO-STEP FLASH PIPELINE ADC

The same principles that lead to the result obtained for the pipeline converter can be applied to the two-step flash architecture. The conceptual block diagram of the well-known two-step flash A/D converter is shown in Figure 7.25. During the quantization decision cycle, the first flash converter quantizes a sampled-and-held input into N_1 digital bits. During the subsequent fine quantization cycle, a reconstructed voltage corresponding to the coarse N_1 bits is subtracted from the sampled-and-held input to generate a residue voltage. This residue voltage is amplified by 2^{N_1} and then applied to the second flash converter to extract the least significant bits.

For the realization of a logarithmic two-step flash A/D converter the operations above have to be carried out in the logarithm domain, where the subtraction is equivalent to a division and the multiplication by 2^{N_1} is equivalent to being raised by 2^{N_1} .

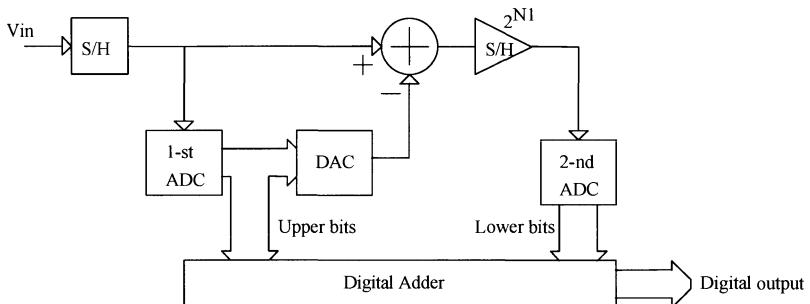


Figure 7.25: Conceptual block diagram of a classical, linear two-step flash ADC.

The flash converters must have a logarithmic quantization law to extract both the most and the least significant bits. The first ADC extracts the most significant bits (in a logarithmic scale), which control the DAC. Then, in the linear case the output of the DAC is subtracted from the input whereas in the logarithmic domain, where the subtraction is equivalent to a division, this operation corresponds to an attenuation that depends on the input digital word from the flash ADC. Based on this approach the DAC should be replaced by a digitally controlled logarithmic attenuator whose attenuation factors a_n are given by

$$a_n = e^{-\frac{\ln\left(1+\frac{1}{K}\right)}{2^N} \cdot 2^{N_2 \cdot n}} \quad \text{for } n = 0, \dots, 2^{N_1} - 1 \quad (7.25)$$

with N_1 and N_2 , respectively, being the resolution bits of the first and second flash converters, and N the total number of bits of the converter. Furthermore, to

use the same flash converter in both the coarse and fine quantization cycles the residue voltage signal is raised to 2^{N_1} .

In order to implement the logarithmic operations described above, we employ the logarithmic two-step flash architecture of Figure 7.26, where the raised operation ensures that the same logarithmic quantization scale can be used in both conversion steps.

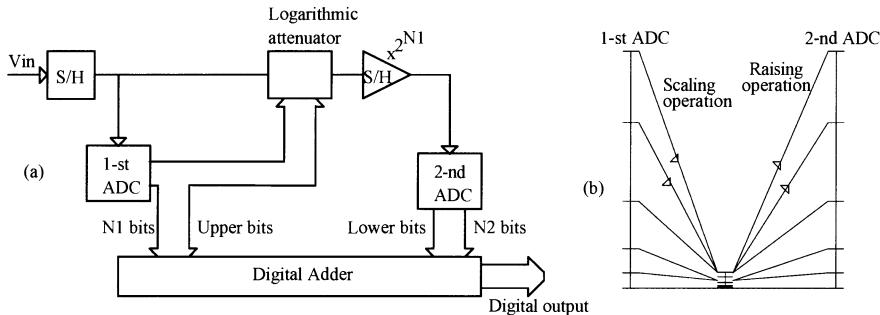


Figure 7.26: (a) Block diagram of a logarithmic two-step flash ADC and (b) signal operation at each step of conversion.

In the basic architecture of Figure 7.26, the realization of the raising function is difficult to implement in integrated circuit form. This, however, is only required to use the same logarithmic scale resolution in both flashes. By abandoning such requirement, and therefore employing a different second flash with the same resolution as for the overall converter, the raising operation is no longer needed and the resulting architecture is similar to a subranging A/D converter [40]. In this case the first and the second ADC's have the threshold voltages respectively given by

$$VT1_n = K \cdot e^{-\frac{\ln(1+\frac{1}{K})}{2^N} \cdot 2^{N_2} \cdot n} \quad \text{for } n = 0, \dots, 2^{N_1}-1 \quad (7.26)$$

$$VT2_n = K \cdot e^{-\frac{\ln(1+\frac{1}{K})}{2^N} \cdot n} \quad \text{for } n = 0, \dots, 2^{N_2}-1 \quad (7.27)$$

However the second ADC must have an offset lower than half LSB of the full converter, given by

$$V_{LSB} = K \cdot e^{-\frac{\ln(1/K+1)}{2^N}} \quad (7.28)$$

Since in a logarithmic scale the threshold voltages depend on their position in the scale, the difference between two adjacent threshold voltages increase as we increase their position. Consequently, a comparator working in a lower position on the scale has to distinguish a voltage lower than its counterpart in a higher position. The lower comparator needs a lower offset and for the same speed will have higher power consumption than its counterpart. This problem can be minimized if we introduce an offset in the scale of the second ADC to allow operation in higher positions of the scale. In the logarithmic domain, where the addition is equivalent to a multiplication, this operation corresponds to an amplification dependent on the scale offset introduced. The higher the offset introduced in the scale, the higher can be the tolerated offset of the comparator. If we shift the second ADC scale near the top of the input range and leave some additional transition steps at the top and at the bottom, it is possible to introduce digital correction of the most significant bits [39]. The precision of the DAC logarithmic attenuator/amplifier can be also calibrated with analog and/or digital techniques.

Figure 7.27 shows the improved architecture with the signal operation at each conversion step and the error correction technique with the amplification step merged in the logarithmic attenuator.

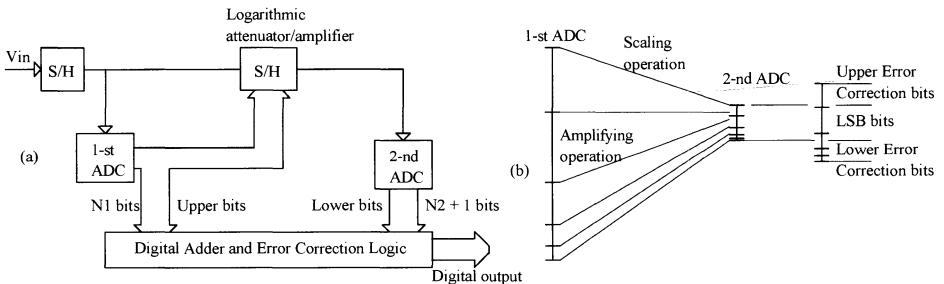


Figure 7.27: (a) Improved architecture of the logarithmic two-step flash ADC and (b) signal operation at each step of conversion.

The resulting computer simulated conversion characteristic of the architecture of Figure 7.27, for 7-bits resolution with $N_1 = 4$ and $N_2 = 3 + 1$, is shown in Figure 7.28. The extra bit in the second ADC is for digital error correction purposes. The amplification that has been introduced will shift all the threshold voltages of the first ADC to the top of the scale and is given by

$$A_n = e^{\frac{\ln(1/K+1)}{2^N} \cdot (2^N - 2^{N_2}(1+n))} \quad \text{for } n = 1, \dots, 2N_1 \quad (7.29)$$

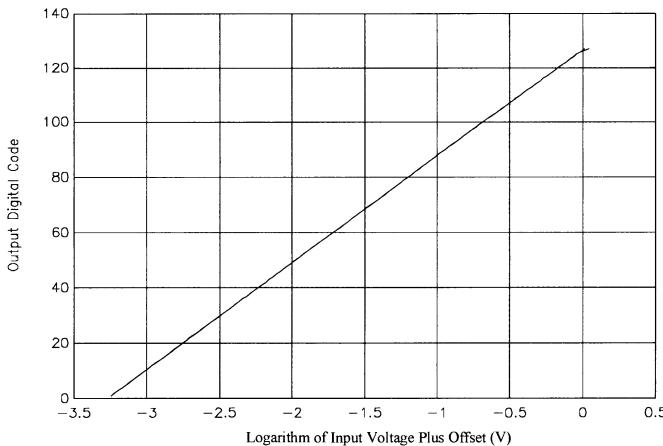


Figure 7.28: MATLAB simulation of the logarithmic two-step flash with 60 dB dynamic range.

Table 7.2 shows the threshold voltages of the first and second ADC and the attenuation/amplification factors for the case of $N_1 = 4$ and $N_2 = 3$ bits and $K = 0.038$ (60 dB dynamic range), corresponding to a total converter resolution of 7 bits.

Table 7.2: Attenuation factors of the digitally controlled logarithmic attenuator.

Comp. stage	Threshold Voltage of the first ADC (V)	Threshold Voltage of the second ADC (V)	Attenuation/amplification factor	Comment (second ADC bits)
1	0.0467	0.6353	18.0660	Lower Correction bit
2	0.0575	0.6519	14.6922	
3	0.0706	0.6690	11.9484	
4	0.0869	0.6865	9.7171	Lower Correction bit
5	0.1068	0.7045	7.9024	LSB bits
6	0.1314	0.7229	6.4266	
7	0.1615	0.7418	5.2265	
8	0.1986	0.7613	4.2504	
9	0.2442	0.7812	3.4566	
10	0.3003	0.8016	2.8111	
11	0.3692	0.8226	2.2861	
12	0.4540	0.8442	1.8592	LSB bits
13	0.5583	0.8663	1.5120	Upper Correction bit
14	0.6865	0.8889	1.2296	
15	0.8442	0.9122	1.0000	
16	1.0380	0.9361	0.8133	Upper Correction bit

Figure 7.29 shows the conversion characteristics and absolute error of the converter of Figure 7.27 with 7-bit resolution, for V_{in} varying from K to $(1 + K)$,

and assuming a precision of 5% and 0.1% for the resistor string values and attenuation/amplification coefficients, and 60 dB gain comparators. Figure 7.30 shows the INL and DNL characteristics of the same converter, with and without digital error correction.

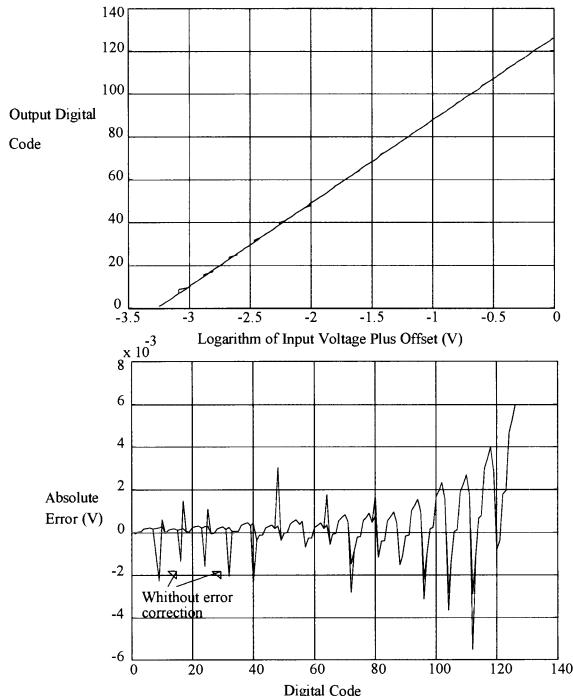


Figure 7.29: Conversion characteristics and absolute error of the logarithmic ADC of Figure 7.28 for 7 bits resolution, 5 % precision on the resistor string, 60 dB gain comparators and 60 dB dynamic range.

The threshold voltages of the first and second ADC are obtained from a resistor string. The value of each one is given by (7.2), (7.3) [24]. In such converter the input voltage range is between K and $1 + K$ (1 V full scale). Should the input voltage be referred to ground, then it is necessary to introduce the offset K at the input as shown in Figure 7.31.

This section shows the principle of logarithmic transformation applied to the case of a two-stage flash. The same principle can also be applied to algorithmic ADC as proposed in [24]. Other possible applications of this transformation are in logarithmic DACs.

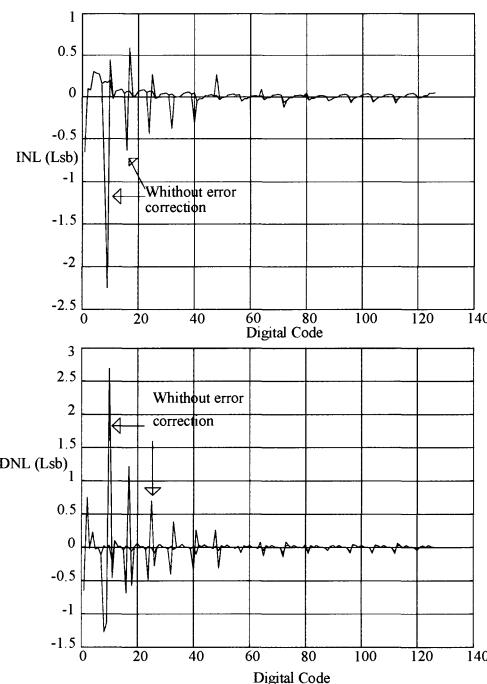


Figure 7.30: INL and DNL of the logarithmic ADC of Figure 7.28 for 7 bits resolution, 5% precision on the resistor string, 60 dB gain comparators and 60 dB dynamic range.

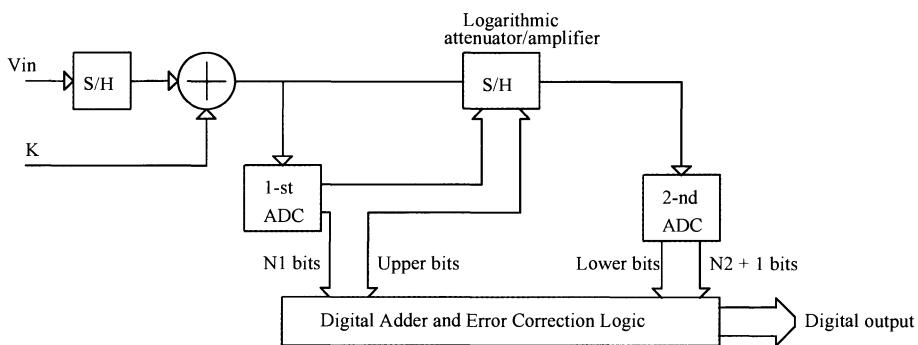


Figure 7.31: Block diagram of the logarithmic two-step flash ADC, with the input voltage referred to ground.

7.9 A LOGARITHMIC PIPELINE ADC WITH 80 dB DYNAMIC RANGE

To illustrate the principles described in the last sections, an integrated prototype has been designed and fabricated. The pipeline architecture is the one shown in Figure 7.19. The summary of specifications of the ADC is shown in Table 7.3, and Figure 7.32 shows the chip layout.

Table 7.3: Summary of the logarithmic 9-bit pipeline ADC

Resolution	9 bit
Conversion rate	10 Ms/s
Power consumption	345 mW
Power supply	2.2 - 2.7 V
Input range	2 Vpp
SNR	45.67 dB
Dynamic range	79.28 dB
Chip area	2.6 mm x 3.1 mm
Process technology	0.25 μ m CMOS single poly, 5-metal, MiM

An experimental result from the logarithmic ADC prototype is shown in Figure 7.33 for a positive triangular waveform. The corresponding output is a direct logarithm of the input. Figure 7.34 and Figure 7.35 shows the correspondent DNL and INL of the logarithmic converter. The obtained INL is limited by the offset of the squarer block that introduce distortion in the transfer characteristic.

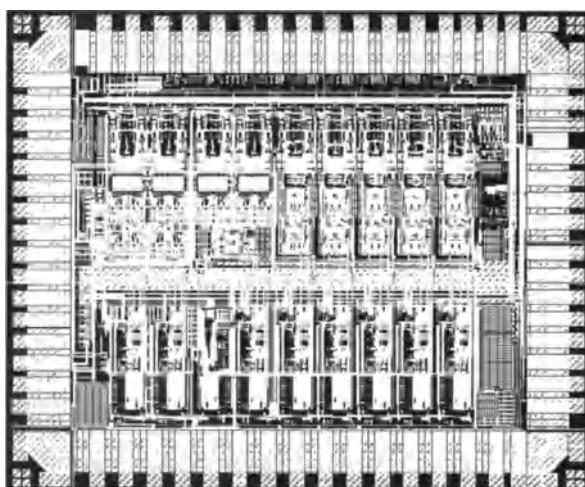


Figure 7.32: 9-bit Logarithmic pipeline ADC layout.

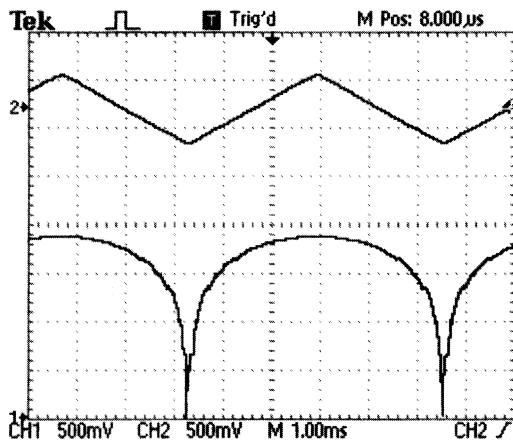


Figure 7.33: Output of the pipeline converter by injecting a positive triangular waveform of 250 Hz at the input, for a clock frequency of 1.6 MHz.

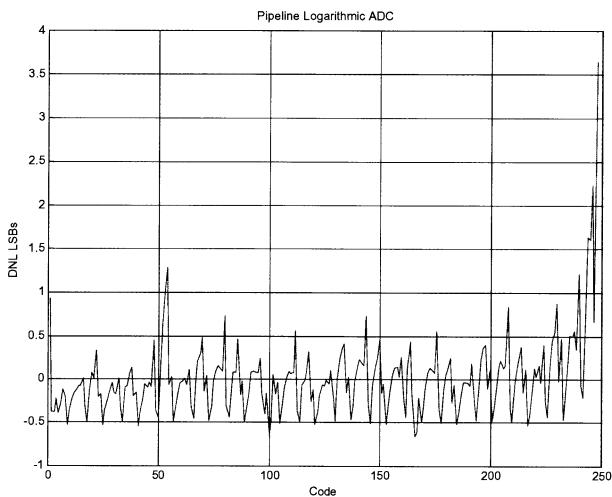


Figure 7.34: DNL of the pipeline converter for a clock frequency of 1.6 MHz.

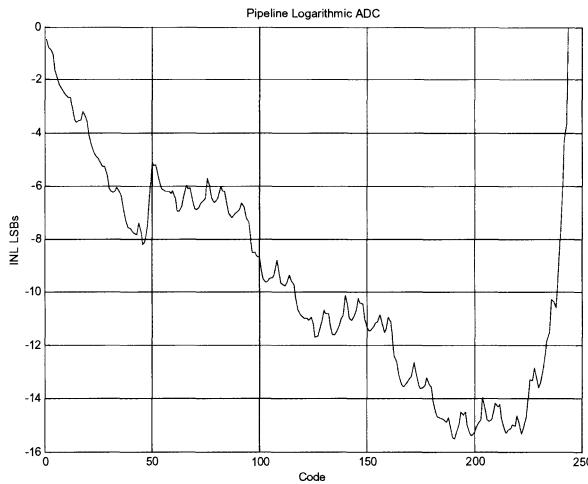


Figure 7.35: INL of the pipeline converter for a clock frequency of 1.6 MHz.

7.10 SUMMARY

This chapter described some of the architectures used in non-uniform quantizers and reviewed some of the most important topologies. A transformation from the linear domain to the logarithmic domain has been proposed. This transformation allows the conversion of a given linear converter topology to the logarithmic domain, by changing the linear signal operations into their equivalent in the logarithmic domain.

This principle was applied to a 1.5bit/stage pipeline ADC, thus transforming it in a logarithmic analog-to-digital converter. Some changes were done in the resulting topology to simplify the circuit implementation and also to improve noise performance of the system. Digital error correction and digital calibration similar to the ones used by linear converters were applied to achieve high resolution and high dynamic range operation. The same operations were also applied to a two-step flash ADC to obtain a logarithmic converter. The two-step flash also employs the digital error correction technique to relax the comparator requirements.

It was shown that the pipeline converter could be used for the targeted ADSL application with the same dynamic range and SNR as the linear pipeline converter with a reduction on the number of bits. Simulations show the performance of the proposed converter.

To be possible to have similar performance metrics in terms of INL and DNL, the same type of logarithmic transformation was done to the converters input

output characteristics. This was needed because there is a lack of performance metrics in the literature for these types of converters. With this transformation the converter transition step is made constant and, in this way, it is possible to measure the INL and DNL with the same definitions that are used for linear converters.

An experimental prototype of a 9-bit 80 dB dynamic range ADC has been integrated. Preliminary evaluation results are presented.

7.11. REFERENCES

- [1] Bruce Carlson, *Communication Systems*, McGraw-Hill Company 1986.
- [2] Transmission performance characteristic of pulse code modulation channels, ITU-T Recommendation G.712.
- [3] Y. P Tsividis, V. Gopinathan and L. Tóth, "Companding in Signal Processing", *Electronic Letters* Vol. 26, no. 17, pp. 1331-1332, 16th August 1990.
- [4] A. K. Oki, M. E. Kim, G. M. Gorman and J. B. Camou, "High-Performance GaAs Heterojunction Bipolar Transistor Monolithic Logarithmic IF Amplifiers", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 36, No. 12, pp. 1958-1965, December 1988.
- [5] Willian L. Barber and Edmund R. Brown, "A True logarithmic Amplifier for Radar IF Applications", *IEEE Journal of Solid State Circuits*, Vol. SC-15, no. 3, pp.291-295, June 1980.
- [6] Katsuji Kimura, "A CMOS Logarithmic IF Amplifier with Unbalanced Source-Coupled Pairs", *IEEE Journal of Solid State Circuits*, Vol.28, No. 1, pp.78-83, January 1993.
- [7] Bill Schweber, "Log and Limiting Amps Tame Rowdy Communications Signals", *Electronic Design News*, pp.117-126, August 1999.
- [8] Douglas R. Frey, "Exponential State Space Filters: A Generic current Mode design Strategy", *IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications*, Vol. 43, no. 1, pp. 34-42, January 1996.
- [9] Gordon W. Roberts and Vincent W. Leung, *Design and Analysis of Integrated-Based Log-Domain Filter Circuits*, Kluwer Academic Publishers, 2000.
- [10] M. N. El-Gamal and G. W. Roberts, "Very High-Frequency Log-domain Bandpass Filters", *IEEE Transactions on Circuits and Systems II*, Vol. 45, No. 9, pp.1188-1198, September 1998.
- [11] S. Cantarano and G. V. Pallottino, "Logarithmic Analog-to-Digital Converters: A Survey", *IEEE Transactions on Instrumentation and Measurement*, Vol. IM-22, No. 3, pp. 201-213, September 1973.
- [12] Henry O. Kunz, "Exponential D/A Converter with a Dynamic Range of Eight Decades", *IEEE Transactions on Circuits and Systems*, Vol. CAS-25, no. 7, pp. 522-526, July 1978.
- [13] Yannis P. Tsividis, Paul R. Gray, David A Hodges and Jacob Chacko, "A Segmented μ-255 Law PCM Voice Encoder Utilizing NMOS Technology", *IEEE Journal of Solid State Circuits*, SC-11, pp. 740-747, December 1976.

- [14] Jorge Guilherme and José E. Franca, "Digitally-Controlled Analogue Signal Processing and Conversion Techniques Employing a Logarithmic Building Block", *Proc. IEEE International Symposium on Circuits and Systems* Vol. 5 pp. 377-380, London May 1994.
- [15] Gottschalk, "Logarithmic Analog-to-Digital Converter using Switched Attenuators", *Rev. Sci. Instrum.* No. 49, pp. 200-204, February 1978.
- [16] Jorge Guilherme, João Vital and José Franca, "New CMOS Logarithmic Two-Step Flash A/D Converts with Digital Error Correction", *Proc. IEEE 38th Midwest Symposium on Circuits and Systems*, pp.881-884, August 1995.
- [17] Graeme J., "Digital gain control streamlines signal-aquisition systems", *Electronic Design News*, pp. 171-180, May 1986.
- [18] Burrier R., et all, "Floating-point converter uses hardware to get a 20-bit dynamic range", *Electronic Design*, pp. 175-186, September 1984.
- [19] J. N Lygouras, "Nonlinear ADC with Digitally Selectable Quantizing Characteristic" *IEEE Transaction on Nuclear Science*, Vol. 35, no. 5, pp. 1088-1091, October 1988.
- [20] Dineh K. Anvekar and B. S. Sonde, "A Programmable Nonlinear ADC Using Optimal-Sized ROM", *IEEE Transaction on Instrumentation and Measurement*, Vol. 40, No. 6, pp.1031-1035, December 1991.
- [21] Zhang Z. and G. S. Temes, "Multi-bit Oversampled $\Sigma\Delta$ A/D Converter with Nonuniform Quantization", *Electronic Letters*, Vol. 27, No. 6, pp. 528-529, March 1991.
- [22] L. Hernandez and S. Paton, "Noise Shaping Modulator with Logarithmic Response", *Electronic Letters*, Vol. 35, No. 12, pp. 955-956, June 1999.
- [23] J. N. Lygouras, "Non-Linear Analog-to-Digital Conversion Through PWM", *International Journal of Electronics*, Vol. 62, No. 3, pp. 347-352, 1987.
- [24] Jorge Guilherme and José Franca, "New CMOS Logarithmic A/D Converters Employing Pipeline and Algorithmic Architectures", *Proc. IEEE International Symposium on Circuits and Systems*, pp.529-532, May 1995.
- [25] Jorge Guilherme, "Circuitos Integrados CMOS para Processamento em Radiocomunicações", *Master Thesis*, Instituto Superior Técnico, Technical University of Lisbon, November 1994.
- [26] B. Hallgren and H. Verweij, "New Developments in Time Pulse Height Digitizers", *IEEE Transactions on Nuclear Science*, Vol. NS-27, No. 1, pp. 333-337, February 1980.
- [27] Dwight U. Thompson and Bruce A. Wooly, "A 15-bit Pipelined CMOS Floating-Point A/D Converter", *IEEE Journal of Solid State Circuits*, Vol. 36, No. 2, pp. 299-303, February 2001.
- [28] Johan Piper and Jiren Yuan, "Realization of a Floating-Point A/D Converter", *IEEE International Symposium on Circuits and Systems*, Vol. I, pp. 404-407, Sydney May 2001.
- [29] Jiren Yuan and Johan Piper, "Floating-Point Analog-to-Digital Converter", *International Conference on Electronics, Circuits and Systems*, Vol. 3, pp. 1385-1388, September 1999.

- [30] F. Francesconi and F. Maloberti, "A Low Power Logarithmic A/D Converter", *IEEE International Symposium on Circuits and Systems*, Vol. I, pp. 473-476, 1996.
- [31] Jorge Guilherme, Nuno Horta and José Franca, "Symbolic Synthesis of Non-Linear Data Converters", *IEEE International Conference on Electronics, Circuits and Systems*, pp. 3.219-3.222, Lisbon September 1998.
- [32] Jorge Guilherme, João Vital and José Franca, "A True Logarithmic Analog-to-Digital Pipeline Converter with 1.5bit/stage and Digital Correction", *IEEE International Conference on Electronics, Circuits and Systems*, pp. 393-396, September Malta 2001.
- [33] Jorge Guilherme, João Vital and José Franca, "A CMOS Logarithmic Pipeline A/D Converter with a Dynamic Range of 80 dB", *IEEE International Conference on Electronics, Circuits and Systems*, Dubrovnik, Croatia, September 2002.
- [34] Thomas F. Schubert Jr., "Simulation of the Effect of Companding on Quantization Noise in Digital Communication Systems", *IEEE Transaction on Education*, Vol. 37, no. 1, pp. 47-50, February 1994.
- [35] Denis J. G. Mestdagh and Lennart B. Olsson, "Transmitter & Receiver Clipping Mitigation in DMT Transceivers", *Analog Circuit Design Series*, Kluwer Academic Publishers 1999.
- [36] IEEE Std 1241, IEEE standard for terminology and test methods for analog-to-digital converters.
- [37] Giovanni Bucci, Marco Faccio and Carmine Landi, "The performance test of a piece-linear A/D converter", *IEEE Instrumentation and Measurement Technology Conference*, pp. 1223.1228, St. Paul USA May 1998.
- [38] Jorge Guilherme, João Vital and José Franca, "Performance Testing of Logarithmic Analog-to-Digital Converters", *11th IMEKO TC-4 Symposium trends in Electrical Measurement and Instrumentation and 6th EuroWorkshop on ADC modeling and testing*, Lisbon September 2001.
- [39] Seung-Hoon Lee and Bang-Sub Song, "Digital-Domain Calibration of Multistep Analog-to-Digital Converters", *IEEE Journal of Solid State Circuits*, Vol. 27, No. 12, December 1992.
- [40] Andrew G. F. Dingwall and Victor Zazzu, "An 8-MHz CMOS Subranging 8-bit A/D Converter", *IEEE Journal of Solid State Circuits*, Vol. SC-20, No.6, pp. 1138-1143, December 1985.
- [41] Carlos Salema, *Feixes Hertesianos*, AEIST, Instituto Superior Técnico, Technical University of Lisbon, Portugal, 1995.
- [42] Jonh Bellamy, *Digital Telephony*, Wiley 1982.
- [43] Rudy Van de Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters*, Kluwer 1994.

Chapter 8

SINGLE-LOOP MULTI-BIT SIGMA-DELTA MODULATORS

Yves Geerts, Michiel Steyaert, Willy Sansen

Abstract $\Delta\Sigma$ converters are suitable to implement high-performance analog-to-digital converters. Several topologies are first reviewed in the context of high-resolution high-speed design targets. The advantages of multi-bit topologies are revealed and several solutions for the linearity requirements of the feedback DAC are discussed. The remainder of this chapter focuses on the influence of several important circuit non-idealities which can become performance limiting factors. A 16-bit 2.5 MS/s converter is discussed as a design example.

8.1 Introduction

Over the last decade, a vast evolution of communication systems was observed driven by the broadband Internet access demands and the development of wireless communication systems. The core of these complex electronic systems consists of digital circuits which have a huge computational power and are implemented in CMOS technologies. These systems require a high-performance interface to the analog world.

This chapter discusses design aspects of high-performance $\Delta\Sigma$ AD converters, with a special focus towards single-loop multi-bit implementations. The first section discusses different architecture options and takes a closer look at the advantages and problems of multi-bit implementations. Section 8.3 takes a look at the optimal implementation of the integrator, the DAC and the sizing of the sampling capacitances. Section 8.4 discusses the influence of several circuit non-idealities on the performance of the converter. Finally, the design of a 16-bit $\Delta\Sigma$ converter is discussed as an example, followed by some conclusions.

8.2 Architectures

In order to compare the performance of different architectures, a reference $\Delta\Sigma$ converter is defined which has an ideal n^{th} -order low-pass noise shaping function represented by $H_e(z) = (1 - z^{-1})^n$, and a simple n^{th} -order delay as the signal transfer function. Under the assumption that the quantization-noise can be represented by an additive white noise source [Raz 95, Can 96], the peak signal-to-noise ratio of the reference structure can be calculated as [Gee 02]

$$SNR_p = \frac{3\pi}{2} \cdot (2^B - 1)^2 \cdot (2n + 1) \cdot \left(\frac{OSR}{\pi} \right)^{2n+1} \quad (8.1)$$

This formula shows the three main parameters which can increase the performance of the converter: the order of the loop-filter (n), the number of bits in the quantizer (B) or the oversampling ratio (OSR).

In order to combine a high resolution with a high speed, a small oversampling ratio should be selected to limit the clock speed of the converter and thereby the bandwidth requirements of the integrators. Therefore, a topology is required which can deliver high accuracies at low oversampling ratios.

Classical single-loop single-bit $\Delta\Sigma$ converters would require a high order of the loop filter to achieve a good accuracy at a low oversampling ratio. Unfortunately, increasing the order seriously degrades the stability, resulting in a serious deterioration of the SNR compared to an ideal n -th order structure [Nor 96, Mar 98b]. This will be illustrated further on.

Cascaded or MASH topologies [Nor 96, Bra 91, Med 99a, Med 99b, Mar 98a, Gee 99] can combine high order noise shaping with the excellent stability of a second order converter. The main drawback of these topologies are the high building-block specifications in order to avoid noise leakage from the first stage to the output, as illustrated in Section 8.4.

Finally, multi-bit converters can achieve a significant improvement in performance by employing a multi-bit quantizer. Besides the accuracy improvement for each extra bit in the quantizer, they offer improved stability. This allows a more aggressive noise shaping which results in an additional accuracy improvement [Nor 96, Mar 98b, Ada 86].

8.2.1 Multi-Bit Single-Loop Topologies

It was shown in (8.1) that increasing the order of the loop-filter improves the accuracy of the converter since the quantization-noise is more suppressed. However, as the order of the loop-filter increases, it becomes

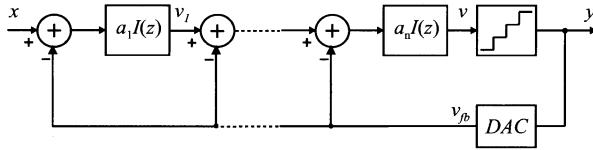


Figure 8.1: Generalized n^{th} -order single-loop multi-bit $\Delta\Sigma$ converter with distributed feedback. Without loss of generality, the coefficients of the feedback branches are set to unity.

more difficult to obtain a stable $\Delta\Sigma$ converter and the gain of the loop-filter has to be reduced, resulting in significant performance loss. This will be illustrated for second and third-order loops.

Fig. 8.1 shows a general representation of a n^{th} -order single-loop $\Delta\Sigma$ with distributed feedback. Without loss of generality, all the feedback branches are shown with a gain of unity as this results in the minimum number of independent parameters. If the quantizer is modeled by an additive white noise-source and a gain k , the signal-to-noise ratio of this structure can be calculated as

$$SNR_p = SNR_{ideal} \cdot \left(k \cdot \prod_{i=1}^n a_i \right)^2 \quad (8.2)$$

where SNR_{ideal} is expressed by (8.1) [Mar 98b]. For multi-bit $\Delta\Sigma$ converters, the quantizer gain k of this model is a well determined number, but for single-bit quantizers it can be chosen arbitrary since the quantizer has only two levels. Since a single-bit quantizer is only sensitive to the polarity of the input signal, the gain of the last integrator a_n is irrelevant to the operation of a single-bit $\Delta\Sigma$ converter. Therefore, the product of the quantizer gain k and the coefficient of the last integrator a_n is combined into $k_{eff} = k \cdot a_n$. In [Mar 98b], a relative good fitting between simulations and the analytical model of single-bit $\Delta\Sigma$ converters is found for $k_{eff} = 2$. Therefore, based on the linear model of the quantizer with a fixed gain, equation (8.2) shows that the SNR of the converter can be improved by increasing the loop coefficients a_i since this results in a better suppression of the quantization error.

However, this model does not give any information concerning the stability of the converter. Therefore, the model presented in [Bai 94] can be used to get insight in the stability of the converter. This model represents the quantizer as a variable quantizer gain $k_i = y/v$, which is dependent on the input signal of the quantizer v . Note that this model is no longer based on the additive white noise approximation since the concept of quantization error is not used. Using this model, a root-locus plot can be generated where

Table 8.1: Topology parameters and achievable performance for second-order single-loop $\Delta\Sigma$ converters with one to four bits in the quantizer. The numbers between brackets indicate the performance relative to an ideal converter, as defined by (8.1).

B	Par (a_1, a_2)	OSR	8	16	32	64	128
1	(0.5,0.5)	$SNR_{0.25}$	20 (-1)	35 (-2)	49 (-3)	66 (-1)	82 (0)
		SNR_p	27 (-7)	42 (-7)	57 (-7)	71 (-8)	86 (-8)
		OL	0.81	0.84	0.86	0.82	0.81
2	(0.6,1.1)	$SNR_{0.25}$	31 (0)	44 (-2)	59 (-2)	74 (-2)	90 (-1)
		SNR_p	41 (-2)	56 (-2)	72 (-1)	86 (-2)	102 (-1)
		OL	0.89	0.87	0.88	0.86	0.80
3	(0.6,1.4)	$SNR_{0.25}$	39 (+1)	55 (+2)	70 (+2)	85 (+1)	99 (0)
		SNR_p	51 (+1)	65 (0)	80 (-1)	96 (0)	112 (+1)
		OL	0.95	0.95	0.93	0.93	0.91
4	(0.7,1.6)	$SNR_{0.25}$	50 (+5)	64 (+4)	79 (+4)	92 (+2)	110 (+5)
		SNR_p	60 (+3)	76 (+4)	90 (+3)	105 (+3)	121 (+4)
		OL	0.96	0.96	0.97	0.96	0.95

Table 8.2: Topology parameters and achievable performance for third-order single-loop $\Delta\Sigma$ converters with one to four bits in the quantizer. The numbers between brackets indicate the performance relative to an ideal converter, as defined by (8.1).

B	Par (a_1, a_2, a_3)	OSR	8	16	32	64	128
1	(0.3,0.4,0.5)	$SNR_{0.25}$	13 (-18)	38 (-14)	60 (-13)	81 (-13)	102 (-13)
		SNR_p	17 (-26)	42 (-22)	65 (-20)	86 (-20)	107 (-20)
		OL	0.56	0.42	0.44	0.43	0.44
2	(0.3,0.5,1.4)	$SNR_{0.25}$	29 (-12)	52 (-10)	73 (-10)	95 (-9)	116 (-9)
		SNR_p	39 (-14)	61 (-13)	83 (-12)	103 (-13)	125 (-12)
		OL	0.84	0.84	0.79	0.80	0.78
3	(0.3,0.6,2.0)	$SNR_{0.25}$	39 (-9)	61 (-8)	82 (-8)	104 (-7)	125 (-7)
		SNR_p	51 (-9)	72 (-9)	93 (-9)	115 (-8)	136 (-8)
		OL	0.92	0.92	0.90	0.89	0.88
4	(0.3,0.7,2.0)	$SNR_{0.25}$	47 (-8)	68 (-8)	89 (-8)	111 (-7)	131 (-7)
		SNR_p	59 (-8)	80 (-8)	102 (-7)	122 (-8)	144 (-7)
		OL	0.97	0.97	0.96	0.95	0.95

the poles move along the locus as the instantaneous gain of the quantizer k_i varies. This method allows to obtain useful information on the stability of the system and clearly illustrates how stable and unstable limit cycles can occur.

In [Gee 02], the methods discussed above are combined with extensive behavioral simulations to obtain insight in the stability issues of $\Delta\Sigma$ converters and to get optimal coefficients for the loop-filters of a large variety of architectures. The results obtained in this study are now briefly discussed, with special attention for the advantages of multi-bit quantizers.

Table 8.1 shows optimal coefficients for a second-order $\Delta\Sigma$ converter with one to four bits in the quantizer and feedback DACs. $\text{SNR}_{0.25}$ represents the SNR for an input amplitude of 0.25 times the reference voltage of the $\Delta\Sigma$ converter and the numbers between brackets indicate the performance degradation compared to the reference structure defined by (8.1). The results show that the degradation compared to the reference structure decreases significantly as the number of quantization bits is increased. This is due to the more aggressive noise-shaping functions in the multi-bit converters since the loop coefficients are increased. For a four-bit quantizer, the performance is even better than for the ideal second-order noise-shaping function. The overload level increases from 0.81 to 0.95 due to the improved stability of the multi-bit $\Delta\Sigma$ converters. Note that a second-order converter can be made intrinsically stable, it will be stable as long as the input is smaller than the overload level and will always recover from overload conditions.

Table 8.2 shows the results for third-order $\Delta\Sigma$ converters. In contrast to first and second-order $\Delta\Sigma$ converters, there is no set of loop-coefficients which can eliminate the possibility of instability, but the boundaries of instability can be influenced by choosing appropriate loop-coefficients. Several options to deal with this problem, such as the use of clipping levels at the output of the integrators, are discussed in [OpT 93, Ada 96]. The results in Table 8.2 show that a single-bit third-order converter performs roughly 20dB worse than the reference structure. This huge performance drop is due to the small loop coefficients which are required to improve the stability of the converter. As the number of bits of the quantizer increase, the quantizer becomes more linear and has a stabilizing effect on the $\Delta\Sigma$ converter. This allows the use of a more aggressive noise-shaping function by increasing the loop-coefficients. For example, the accuracy of a four-bit third-order design improves by approximately 37dB compared to a single-bit design, of which 13.5dB is due to the more aggressive noise-shaping function and the larger overload level of the converter. This shows that the stabilizing effect of multi-bit converters can be used to gain a significant performance improvement.

For a single-bit fourth-order $\Delta\Sigma$ converter, optimal coefficients are presented in [Mar 98b] and [OpT 93], showing a degradation of 30 to 40dB compared to an ideal fourth-order noise-shaping function.

8.2.2 Performance Comparison

The performance of several $\Delta\Sigma$ topologies, as derived in [Gee 02], is now compared. Fig. 8.2 shows the maximum achievable SNR_p versus the over-

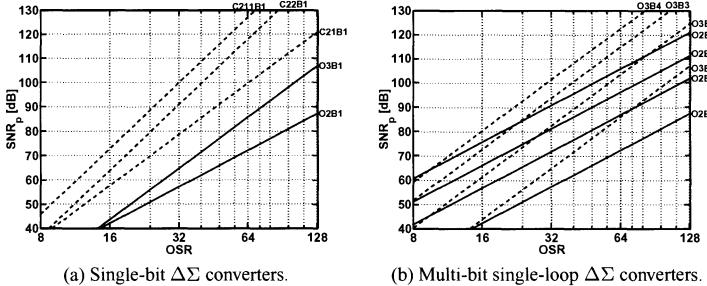


Figure 8.2: Maximum achievable SNR_p versus the oversampling ratio for different $\Delta\Sigma$ topologies. The different topologies are denoted as follows. O_{nBm} is an n^{th} order single-loop converter with a m -bit quantizer and $C_{ijk}B_m$ is a cascaded topology consisting of an $i^{\text{th}}, j^{\text{th}}$ and k^{th} -order stage with m -bit quantizers.

sampling ratio for different topologies. Fig. 8.2a compares the performance of several single-bit $\Delta\Sigma$ converters. Although a third-order single-loop $\Delta\Sigma$ converter performs better than a second-order, the increase of performance is much less than expected from (8.1). This can be explained by the use of smaller loop-coefficients due to the stability issues in third-order converters [Mar 98b].

When comparing the single-loop and cascaded third-order converters, the latter increases the accuracy by more than two bits. This illustrates that cascaded converters can achieve high-order noise-shaping without sacrificing performance due to stability problems. The cascaded C22 and C21 both offer fourth-order noise-shaping, but the C211 outperforms the C22 by approximately one bit.

Clearly, the cascaded 2-1-1 topology offers the best performance of the single-bit converters. However, due to the effect of noise-leakage, cascaded converters require more demanding building block specifications [Rib 91, Yin 93]. This generally leads to an increase of the power consumption and/or circuit complexity. Therefore, depending on the required specifications, a single-loop converter with a higher oversampling ratio may provide a good alternative to a cascaded topology with a lower oversampling ratio.

Fig. 8.2b compares multi-bit single-loop converters. This graph shows the huge increase in performance compared to single-bit designs. Comparing a third-order four-bit converter to a one-bit implementation, a performance increase of 37dB can be observed. Three extra bits in the quantizer are expected to improve the accuracy by 23.5dB due to the decrease of the step-size of the quantizer, which leads to a lower quantization error. However, due to the stabilizing action of the multi-bit quantizer, a more aggressive

noise-shaping function can be used resulting in a better suppression of the quantization noise. Furthermore, also the overload level is increased, leading to a larger signal-power. These last two effects result in an extra improvement of 13.5dB. The performance of the third-order four-bit $\Delta\Sigma$ converter is only one bit less than expected by (8.1) for the reference $\Delta\Sigma$ converter. The same effect can be observed in the second-order $\Delta\Sigma$ converters. A second-order four-bit $\Delta\Sigma$ converter even performs slightly better than expected from (8.1).

Comparing the high-performance cascaded single-bit $\Delta\Sigma$ converters with the single-loop multi-bit converters, shows that the latter offer improved performance in the region of low oversampling ratios. The accuracy of high-order topologies only benefits in combination with a large oversampling ratio, but multi-bit quantization offers an intrinsic improvement of the accuracy for all oversampling ratios. For all oversampling ratios smaller than 64, the third-order four-bit converter outperforms the cascaded 2-1-1 topology and is therefore an interesting alternative for the implementation of high-resolution high-speed converters. Remember that the multi-bit single-loop topologies do not suffer from noise-leakage. Therefore, the building block specification will be more relaxed than for a cascaded implementation.

The main problem of multi-bit $\Delta\Sigma$ converters is the linearity requirement imposed on the DAC in the feedback path. This problem and possible solutions are discussed next. As the number of bits is increased, the complexity of the implementation of the quantizer and the DAC also increases.

8.2.3 Linearity Issues of Multi-Bit $\Delta\Sigma$ Converters

In a multi-bit $\Delta\Sigma$ converter, any non-linearity of the DAC in the global feedback loop around the $\Delta\Sigma$ converter will appear directly at the output without being suppressed or shaped by the $\Delta\Sigma$ loop. The errors introduced by this DAC are subtracted from the input signal and cannot be distinguished from it. Therefore, the linearity of the first DAC needs to be roughly as good as the overall linearity of the $\Delta\Sigma$ converter in order not to degrade the performance. The errors of the other DACs are not introduced at the input node, but at a later stage in the $\Delta\Sigma$ loop. Therefore, these errors are suppressed by the gain of the preceding integrators in the loop and the linearity requirements are less stringent. Note that in a one-bit $\Delta\Sigma$ converter this linearity problem is not an issue since a one-bit DAC only contains two different levels and is therefore intrinsically linear.

Different possibilities to relax the specifications of the DAC are now discussed. A first way is the use of calibration techniques. Note that a contin-

uous background calibration is required to compensate for matching variations due to age and temperature and to allow a continuous operation of the converter. The main drawbacks of these techniques are the increased circuit complexity and the larger time required for testing during production [Cat 89, Gro 89, Sar 93, Fat 93, Bai 96, Car 96].

A second option are the dual-quantization topologies [Hai 91, Bra 91, Med 99a, Med 99b]. The basic idea is to combine the reduced quantization noise of a multi-bit quantizer with the intrinsically linear feedback of a single-bit DAC. It uses an inherently linear single-bit feedback to the first integrator of the loop since any non-linearity of this DAC is immediately visible at the output without any suppression. The feedback to the last integrator in the loop is implemented by a multi-bit DAC since the non-linearities of this DAC are suppressed by the gain of the preceding integrators in the loop. Note that this topology is no longer a true multi-bit structure since the feedback to the first stage is now implemented with only one bit. Therefore, the benefits due to the improved stability are partly sacrificed, resulting in a performance loss compared to full multi-bit topologies. Dual-quantization can also be employed in cascaded topologies, where the first stages use single-bit quantizers and DACs, and the later stages use multi-bit $\Delta\Sigma$ loops. Therefore, the single-bit output of the first stage contains much more quantization noise than the subsequent multi-bit stages and the already severe building block specifications of cascaded topologies are increased even further to avoid degradation due to noise leakage.

A different approach to reduce the linearity requirements of the DAC is offered by dynamic element matching (DEM). In a traditional DAC without DEM, each bit of the thermometer code controls one specific unit element of the DAC. Due to process-variations, the values of these unit elements will not be equal and the DAC will introduce errors and a one-to-one correspondence exists between the input-code of the DAC and the error that it introduces. When dynamic element matching is used, this one-to-one correspondence is broken by a digital element selection block. This digital block tries to select the unit elements in such a way that the errors introduced by the DAC average to zero over multiple time instances. The average output now corresponds to the ideal output. In other words, the errors due to component mismatch are moved to higher frequencies. When the DAC is oversampled, the error falls outside the signal band and can be removed by filtering [Bai 95]. It should be stressed that DEM only compensates the variations of the unit elements and cannot compensate for systematic errors,

such as a finite output impedance of the current-sources in a current-steered DAC [Rad 00].

A large variety of DEM techniques exist and extensive comparisons of different algorithms can be found in [Bai 95, Car 96, Gee 02]. Algorithms such as Data Weighted Averaging (DWA) [Bai 95] and several similar algorithms such as bi-directional DWA [Fuj 00] provide a first-order shaping of the DAC error and they can relax the matching requirements of the unity elements by several orders of magnitude. This will be illustrated in the design example in Section 8.5. The main advantage of DEM compared to dual-quantization techniques is that the $\Delta\Sigma$ converter remains a fully multi-bit structure and can benefit from the improved stability of these structures, allowing more aggressive noise-shaping functions and improved performance.

8.3 Implementation of the Integrator

The feedback of the $\Delta\Sigma$ converter can be implemented in different ways. First, the possibilities for a single-bit feedback are discussed, followed by a discussion of multi-bit feedback. The principles are illustrated in a single-ended way, although implementations are generally differential since this improves the $\text{SNR}_{kT/C}$ by 3dB and provides better suppression of even order harmonics.

There are generally two different ways to implement a single-bit feedback in a switched-capacitor integrator. Fig. 8.3a uses a single reference voltage which is used as the input to both an inverting and a non-inverting network. The decision of the comparator determines how both branches are connected to the differential integrator during clock-phase ϕ_2 . Fig. 8.3b uses two symmetrical reference voltages which can be connected to the input terminal of both sampling capacitances to perform the feedback function.

Under the assumption that C_S equals C_{FB} , both circuits have the same function but some important differences exist with major implications on the power consumption and speed of the integrator.

Since the implementation with double reference voltages shares the sampling capacitance to sample the input and subtract the feedback, it only requires half the number of capacitances and less switches. This results in half the amount of kT/C noise power and consequently the capacitance sizes for the double reference voltage implementation can be reduced by a factor two for the same $\text{SNR}_{kT/C}$. This results in a power decrease of the OTA and a smaller die-size. On top of that, the capacitive feedback factor during the integration phase is larger since less capacitors are connected to the input of the amplifier. This results in a smaller capacitive load and a

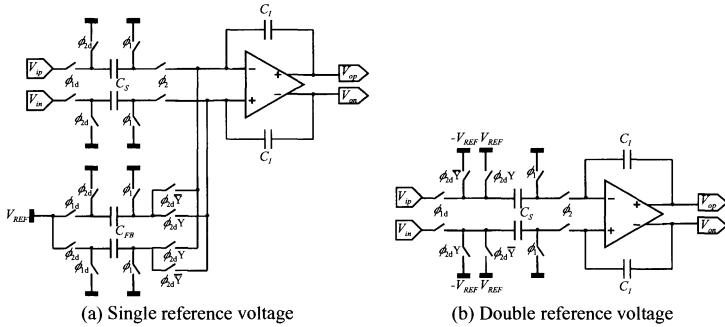


Figure 8.3: Implementation of the feedback with single or double polarity reference voltages. Y represents the outcome of the quantizer.

larger dominant closed-loop pole, leading to a faster settling of the charge transfer.

Unfortunately, there are also a few drawbacks associated with the use of double polarity reference voltages. The circuitry to generate and buffer the reference voltages is more complicated since two symmetrical reference voltages are required. Furthermore, the charge that these buffers deliver to the capacitors is dependent on the input signal. Therefore, special attention must be paid to the reference buffers to ensure sufficient settling performance. Note that the power consumption of the buffer can become an important part of the total power consumption of the $\Delta\Sigma$ converter [Bro 97]. Despite this drawback, double polarity reference voltages are chosen due to the power and settling advantages of the integrators.

When a multi-bit feedback path is required, a first implementation option is to use a resistor ladder between V_{REF} and $-V_{REF}$ to generate the required feedback voltages, as shown in Fig. 8.4a. The switches are driven by a 1-of- n code and connect the wanted ladder taps to the sampling capacitances. Since the resistance ladder is also needed for the multi-bit quantizer, the extra hardware is limited to the switches and a simple digital conversion from the thermometer output code of the quantizer to a 1-of- n code. This hardware can be shared among all the integrators. In order not to degrade the settling performance of the integrators, the resistance of the ladder has to be small, resulting in an increased power consumption. Another drawback of this implementation is the incompatibility with dynamic element matching techniques since these require the shuffling of the unit elements of the ladder.

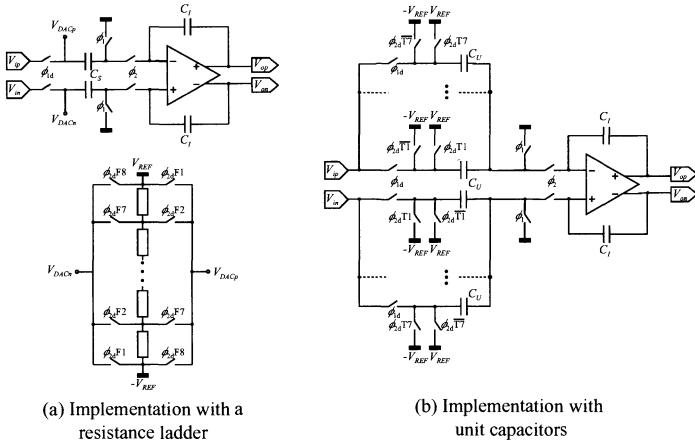


Figure 8.4: Implementations of the multi-bit feedback.

Therefore, this implementation is mainly used in the last stage of dual-quantization $\Delta\Sigma$ converters [Bra 91, Med 99a, Med 99b].

Fig. 8.4b shows a more interesting alternative. For a three-bit implementation, the sampling capacitance is split up into seven parallel unit capacitors, which can be connected separately to either V_{REF} or $-V_{REF}$ during the integration phase. The connection of the unit capacitors to V_{REF} or $-V_{REF}$ is directly controlled by the thermometer output code of the quantizer. Since the use of each unit capacitor can be directly controlled, this implementation is very well suited for dynamic element matching techniques.

To determine the size of the sampling capacitance, two requirements have to be taken into account. The first requirement is related to the circuit noise of the converter. In most practical cases, the noise of the OTAs can be neglected compared to the kT/C noise [Gee 02] and the circuit noise for a differential implementation can be expressed as

$$SNR_{kT/C,\text{approx,diff}} = \frac{(2 \cdot OL \cdot V_{REF})^2}{2} \cdot \frac{OSR \cdot CS}{4kT} \quad (8.3)$$

where OL represent the overload level of the converter and OSR the oversampling ratio. The relative level of the quantization noise and the circuit noise are important for the power consumption of the converter. If the total in-band circuit noise is lower than the in-band quantization noise, power is wasted since the sampling capacitors are oversized. On the other hand, when the circuit noise dominates, a lower order converter or a lower oversampling

ratio could be used to obtain the same performance. These considerations show that the in-band circuit noise should be approximately equal to the in-band quantization noise to obtain a low power consumption [Pel 97]. This condition results in a minimum value for the sampling capacitance $C_{S,kT/C}$.

The second requirement for the sampling capacitance is related to the accuracy of the multi-bit feedback DAC. This accuracy depends on the chosen architecture and the use of dual-quantization or DEM algorithms to relax these specifications. Monte-Carlo simulations, combined with matching data of the technology, yield a minimum size for the unit capacitance and thus also for the sampling capacitance, noted as $C_{S,\sigma}$.

Depending on the specifications of the $\Delta\Sigma$ converter, the topology, the number of bits in the quantizer, the dynamic element matching algorithm and the technology, either the specification due to the matching $C_{S,\sigma}$ or due to the kT/C noise $C_{S,kT/C}$ will dominate. If $C_{U,\sigma} > C_{U,kT/C}$, the total sampling capacitance will be larger than needed for the kT/C noise floor. This means a waste of power simply due to the matching constraints of the multi-bit feedback DAC. This situation should be avoided by a proper selection of the architecture. If $C_{U,\sigma} < C_{U,kT/C}$, the capacitances are determined by the kT/C noise floor and no power is wasted. Thus, when a good architecture choice is made, the multi-bit integrator can be implemented with the same amount of capacitance as a one-bit implementation with the same accuracy goal, thereby avoiding additional kT/C noise and capacitive loading of the integrator. The power consumption remains the same as for a single-bit implementation and the die size is also comparable.

8.4 Circuit non-idealities

Due to the influence of several circuit non-idealities, the performance of a practical implementation of a $\Delta\Sigma$ converter can be significantly worse than the values shown in Fig. 8.2. In order to properly design the $\Delta\Sigma$ converter, one should know the influence of all the non-idealities on the performance of the converter. First, several important non-idealities of the switched-capacitor integrator are discussed, followed by a closer look at the quantizer.

The derivation and the equations of the models for these non-idealities are not discussed in this chapter, but a detailed description can be found in [Gee 99, Gee 02, Gee 00b]. Instead, this chapter focuses on the influences of the architecture choice on the specifications of the building-blocks and presents some general design guidelines and rules-of-thumb for various non-idealities.

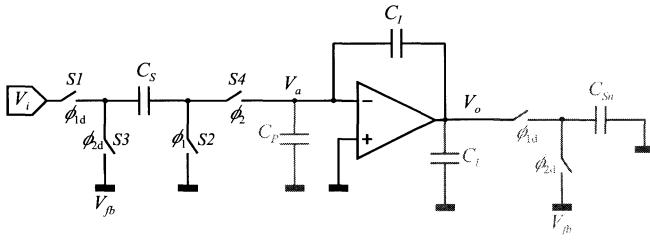


Figure 8.5: Model of a multi-bit switched-capacitor integrator.

The model for the switched-capacitor integrator is shown in Fig. 8.5. V_{fb} represents the feedback of the DAC and C_S and C_I are the sampling and integration capacitances. In order to obtain an accurate model, it is very important to include the parasitic capacitors C_P and C_L associated with the input and output capacitances of the OTA, respectively. C_L also includes the bottom plate parasitic of C_I . C_{Sn} is included to model the sampling operation of the next stage. The model assumes that the OTA of the next integrator is ideal such that the top plate of C_{Sn} is connected to a perfect (virtual) ground during both clock phases.

8.4.1 Finite OTA gain

The finite gain of the OTA is the first non-ideality of the integrator that is discussed. The OTA is represented by a voltage controlled voltage source with gain A . By applying the methods described in [Tem 80], the output of the integrator at the end of the sampling phase can be calculated as [Gee 99]

$$V_o(z) = \frac{C_S}{C_I} \cdot \frac{\rho_2 \cdot [z^{-1}V_i(z) - z^{-1/2}V_{fb}(z)]}{1 - \frac{\rho_2}{\rho_1} z^{-1}} \quad (8.4)$$

where ρ_1 and ρ_2 are the closed-loop static errors and f_{dc1} and f_{dc2} are the capacitive feedback factors during the sampling and the integration phase, respectively. They are given by

$$\rho_1 = \frac{Af_{dc1}}{1 + Af_{dc1}} \quad \rho_2 = \frac{Af_{dc2}}{1 + Af_{dc2}} \quad (8.5)$$

$$f_{dc1} = \frac{C_I}{C_P + C_I} \quad f_{dc2} = \frac{C_I}{C_S + C_P + C_I} \quad (8.6)$$

The finite OTA gain introduces two errors in the transfer function of the integrator. The gain error reduces the gain of the integrator by ρ_2 and

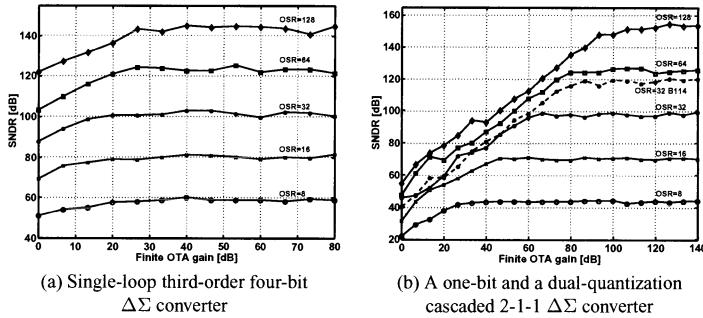


Figure 8.6: Influence of a finite OTA gain on the performance of $\Delta\Sigma$ converters.

the pole error moves the pole from dc ($z=1$) to $z=\rho_2/\rho_1$. Both these errors depend on the product of the OTA gain and the capacitive feedback factors.

This model can be used to determine the required OTA gain. Fig. 8.6 shows simulations of single-loop and cascaded topologies for various oversampling ratios and different OTA gains. The left graph shows the simulation results of a third-order four-bit $\Delta\Sigma$ converter. This shows that a gain of only 30 to 40dB is sufficient in order to avoid performance degradations. Other single-loop topologies have similar characteristics and gain requirements. Note that the gain requirement depends only slightly on the oversampling ratio. This is no longer true when cascaded topologies are considered. The full lines in the right graph show the simulation results of a 2-1-1 cascaded topology with one-bit quantizers. As the oversampling ratio increases, the gain requirement of the OTA increases from 40dB to more than 100dB. This large gain variation shows that the noise-leakage of the quantizer of the first stage to the output of the converter becomes more important as the oversampling ratio increases. This results in tougher building block specifications, such as the OTA-gain. The dashed line in the right graph indicates a dual-quantization cascaded topology with single-bit quantizers in the first two stages and a four-bit quantizer in the last stage. When this curve is compared to the single-bit cascaded implementation with the same oversampling ratio, it is clear that the increased performance of the dual-quantization structure comes at the cost of a severe increase of the gain requirement.

In a practical implementation of an OTA, the gain is not the same for all values of the output voltage. Instead, it decreases as the differential output voltage increases. This is due to the reduction of the output resistance as

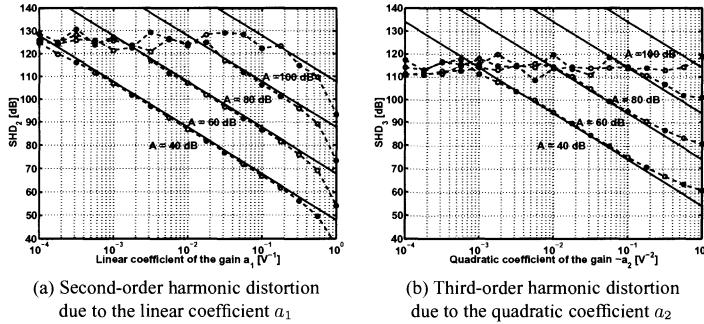


Figure 8.7: Simulation of a third-order four-bit $\Delta\Sigma$ converter with a non-linear OTA gain. The input amplitude A_i equals $0.9V_{REF}$. The full line represents the analytical model, while the dashed lines indicate the simulated values.

the drain-source voltage v_{DS} of the output transistors decreases. The non-linear gain of the OTA can be modeled by a truncated Taylor expansion as

$$A(v) = A_0 \cdot (1 + a_1 v + a_2 v^2) \quad (8.7)$$

where v is the output voltage of the OTA. Note that a_2 is a negative value since the gain decreases as the output swing increases. In [Gee 02], the following signal-to-harmonic-distortion ratios are calculated for this model

$$SHD_2 \approx -20 \log_{10} \left(\frac{a_1}{2A_0} A_i \right) \quad SHD_3 \approx -20 \log_{10} \left(\frac{a_2}{4A_0} A_i^2 \right) \quad (8.8)$$

where A_i is the amplitude of the input signal. These calculations show that the distortion components can be suppressed by increasing the gain of the OTA.

Fig. 8.7 compares behavioral simulations to the analytical model for a third-order four-bit $\Delta\Sigma$ converter for various values of the nominal OTA gain. This shows that the model fits very well, except for very small and very large values of a_1 and a_2 . The deviation for small values is easy to explain. Even when no harmonic distortion components are visible in the output spectrum of the converter, a shaped quantization noise floor is present. This limits the maximum observable SHD_2 and SHD_3 values since the harmonic distortion components are submerged in the quantization noise. Note that the level of the quantization noise in each bin is dependent on the number of points that are used in the simulations. However, when a_1 and $-a_2$ are very close to unity, SHD_2 is up to 12dB worse and SHD_3 up to 6dB better than predicted by the analytical model. At that point, $A(v) \gg 1$ is no longer

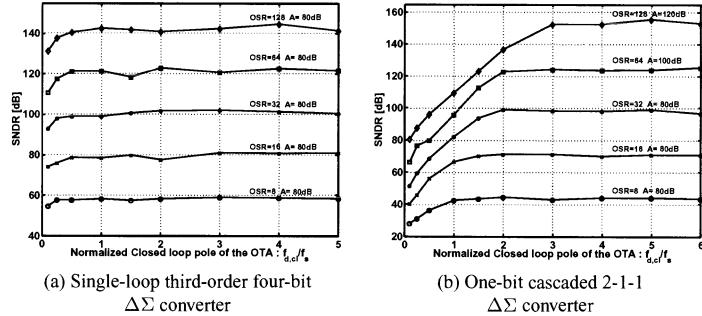


Figure 8.8: Influence of the dominant closed-loop pole of the OTA on the performance of $\Delta\Sigma$ converters.

valid and some approximations made during the derivation of the analytical model are no longer valid.

When a gain of 40dB is selected for the OTA, the gain variation due to a_2 can only be 0.05dB to achieve 16-bit performance levels for a 1V signal. These kind of values are almost impossible to achieve. So although the simulations with a fixed gain only require a gain of 40dB for single-loop topologies, this specification should be increased in order to reduce the distortion components sufficiently.

8.4.2 Finite dominant closed-loop pole of the OTA

In the previous model, all the voltages instantaneously take their final value at the beginning of the sample and integration phase since the voltage controlled voltage source does not model any settling effects. However, in a practical implementation, the poles of the OTA limit the settling performance. Therefore, the OTA is modeled with a transconductance g_m and a finite output conductance g_o to study the influence of the dominant closed-loop pole of the amplifier [Gee 99].

The dominant closed-loop pole during the integration phase can be calculated as

$$p_{cl2} = \frac{g_m}{C_{eq,cl2}} = \frac{g_m}{\frac{C_L + (C_S + C_P)f_{dc2}}{f_{dc2}}} = \frac{g_m}{C_S + C_P + \frac{(C_S + C_P + C_I)C_L}{C_I}} \quad (8.9)$$

where C_L represent the total capacitive load at the output of the OTA and thus also includes the sampling capacitance of the next stage.

Just as for the finite gain of the OTA, simulations are performed for a single-loop and a cascaded $\Delta\Sigma$ converter. Fig. 8.8 shows the simulation

results and indicates the OTA gain which is used during the simulation. An OTA gain of 80dB is chosen for all cases, except for the large oversampling ratios of the cascaded converter. For the single-loop converters, a dominant closed-loop pole equal to 1.5 times the sampling frequency is sufficient. For the cascaded converters, the requirement varies from 1.5 to 4 times the sampling frequency as the oversampling ratio increases. As a rule of thumb, $f_{d,cl} = 1.5 \cdot f_s$ can be chosen for all single-loop topologies. The requirement for cascaded converters depends on the topology and the oversampling ratio, but is generally larger than for single-loop converters.

8.4.3 Resistance of the switches

Up to this point, the switches have been assumed to have an ideal zero resistance when they are closed. In practice, they are implemented with nMOS and/or pMOS transistors exhibiting several non-ideal effects such as a non-zero resistance, clock feedthrough, charge injection and the variation of the resistance with the input signal.

First, the influence of a fixed non-zero resistance is discussed. The resistance of switches S1 and S2 of Fig. 8.5 are lumped into one element with a resistance R_1 and switches S3 and S4 are represented by resistance R_2 . For a single-bit $\Delta\Sigma$ converter, this model introduces no approximations. However, for a multi-bit $\Delta\Sigma$ converter, an exact representation requires a number of parallel branches with a unit capacitance, a resistance and a feedback to $\pm V_{REF}$, corresponding to Fig. 8.4b. This would result in a very complicated high-order model, since the number of nodes drastically increases. Therefore, the approximative model of Fig. 8.5 is used, but the results correspond very well to the more complex model [Gee 02].

Compared to the equation for the settling error in the previous section, which does not include the switch resistance, the dominant closed-loop pole is degraded to

$$p_{cl2,R} = \frac{p_{cl2}}{1 + p_{cl2} \cdot R_2 C_S} \quad (8.10)$$

where p_{cl2} is given by (8.9). This clearly shows the influence of the switch resistance during the integration phase on the degradation of the settling behavior of the integrator.

Fig. 8.9 shows simulations to determine the specifications for the on-resistance of the switches. In order to determine the influence of the resistance during the sampling phase (R_1), during the integration phase (R_2) and the combined effect of both, three simulations are shown which are represented by the dashed, dash-dot and solid lines, respectively. For the

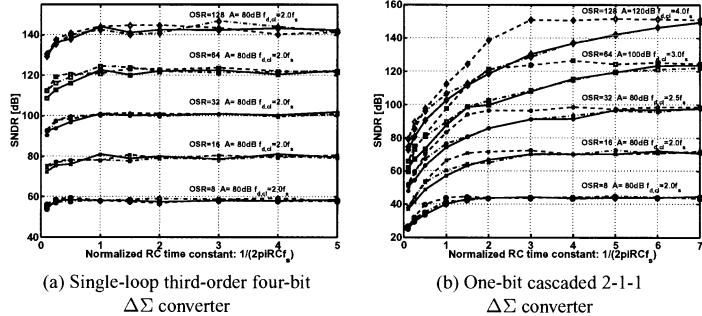


Figure 8.9: Influence of the non-zero resistance of the switches on the performance of $\Delta\Sigma$ converters. The full lines represent simulations with $R_1 = R_2$. The dashed (dash-dot) lines show the influence of R_1 (R_2) while R_2 (R_1) is kept very small.

single-loop $\Delta\Sigma$ converter, all three lines are quite close together and it is sufficient that $\frac{1}{2\pi R C_S}$ is larger than 1.5 times the sampling frequency. For the cascaded 2-1-1 converter with a low oversampling ratio, the same conclusion can be drawn. However, as the oversampling ratio increases, it becomes clear that the resistance during the sampling phase R_1 is of lesser importance. This could be expected from the model since R_2 and C_S not only form a time constant which slows down the charge transfer, but also reduces the dominant closed-loop pole of the OTA. Just as for the previous models, a larger oversampling ratio requires tougher specifications for the cascaded converters. This comes down to a smaller switch resistance and therefore larger switches resulting in more clock feedthrough and increased capacitive loading of the clock buffers that drive these switches. Note that the dominant closed-loop pole influences the specifications of the switch resistance significantly. A larger dominant closed-loop pole generally results in a more relaxed specification for the switches, as can be seen from (8.10).

The second important problem related to the switches is the variation of the switch-resistance with the input signal. When a switch is in the on-phase, it can be assumed that the transistor is in the linear operation region. The resistance of an nMOS is given by [Lak 94]

$$R_N = \frac{1}{K P_n \left(\frac{W}{L}\right)_n \cdot \left((v_G - V_{Tn}) - \frac{v_S + v_D}{2}\right)} \quad (8.11)$$

This equations shows two important effects. First, a reduction of the supply voltage immediately increases the switch resistance since the overdrive voltage of the transistor decreases. Second, the switch resistance is dependent

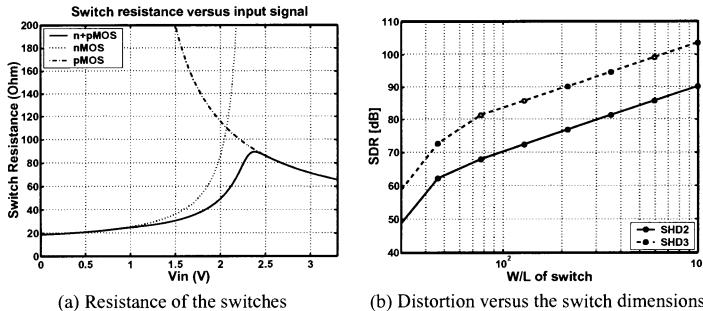


Figure 8.10: Harmonic distortion due to the non-linear resistance of the switches. The switches have a width of $120\mu\text{m}$ and a length of $0.5\mu\text{m}$. The power supply equals 3.3V . The input and sampling frequency are 825kHz and 52.8MHz , respectively. The signal amplitude equals 0.5V .

on the source and drain voltages. When looking at the switched-capacitor integrator of Fig. 8.5, it is obvious that the resistance of switch S1 depends directly on the input signal of the $\Delta\Sigma$ converter. This generates harmonic distortion. Both these effects can be reduced by employing transmission gates with n and pMOS transistors in parallel.

The simulated switch resistance of switch S1 as a function of the input signal is shown in Fig. 8.10a. This graph clearly illustrates that the resistance of switch S1 depends directly on the input signal and thus causes harmonic distortion. Switch S2 always has one terminal connected to a fixed voltage. So at the end of the sampling phase, the voltages at the source and drain of that switch are about constant for each clock period. Therefore, the distortion generated by this switch can be neglected compared to the distortion introduced by switch S1. The same applies to switches S3 and S4, which are connected to respectively a fixed reference voltage or the virtual ground of the OTA at the end of the integration. Furthermore, no time varying input signal is driving the circuit during the integration phase. Instead, a constant charge proportional to the input signal is transferred from the sampling to the integration capacitance. Since this is like applying a dc signal during every clock phase, S3 and S4 generate considerably less distortion. Therefore, to study the distortion introduced by the switches, only the sampling operation through switches S1 and S2 has to be considered.

Fig. 8.10b shows that SHD_2 and SHD_3 improve approximately by $20\text{dB}/\text{dec}$ as the dimensions of the switches increase. For a differential implementation, SHD_2 is sufficiently suppressed, but one should ensure that SHD_3 is larger than the wanted resolution of the converter for an input signal close to

overload and at the edge of the signal band since these conditions ensure worst case distortion components [Gee 99]. This can be done by making the switches large enough. Note that this improves both the distortion and the settling performance of the integrator. A drawback is that larger switches increase the clock feedthrough, the charge injection and the capacitive load on the clock drivers. If the switch becomes too large, the non-linear parasitic junction capacitances of the switch can eventually dominate the sampling capacitance and degrade the linearity of the sampling operation.

As the supply voltage of modern technologies is further reduced, special circuit techniques and technologies can be used to improve the linearity of the input switch.

A first option is the use of technologies with low V_T devices since this leads to a larger overdrive voltage and a reduced on-resistance, but these devices tend to have problems with increased leakage currents. Besides this, the extra processing steps lead to increased cost and turn-around time [Fuj 00, Bul 00].

Another way to improve the linearity of the sampling operation is by applying clock boosting techniques [Bul 00]. A first technique boosts the clock signals to twice the supply voltage to reduce the resistances [Cho 95]. A second technique tries to keep the overdrive voltage V_{GST} of the switches constant by boosting the clock signal to the power supply plus the input signal [Bro 97, Abo 99, Des 01]. The constant V_{GST} significantly enhances the linearity, although the body effect still makes the resistance signal dependent. Two variations of this techniques exist. The first has a constant gate voltage during the sampling process [Bro 97]. This voltage equals the input signal plus the power supply. It does not track any variation of the input signal during the sampling operation. In contrast to this, the gate voltage of the second technique tracks the sum of the input signal and the power supply during the sampling operation and ensures a constant V_{GST} [Abo 99, Des 01]. This last variation is more suitable for input frequencies close to the Nyquist rate. Finally, [Pan 00] proposes a technique which also compensates for the body effect by using a replica. The drawback is the requirement for a high-speed OTA and thus the large power consumption.

All these boosting techniques require the use of voltage levels above the intended supply voltage of the technology. Although some of these techniques ensure that V_{GS} and V_{DS} are always below the maximum power supply [Abo 99, Des 01], care must be taken not to compromise the lifetime of the circuits. Another drawback is the area and power overhead of these boosting circuits.

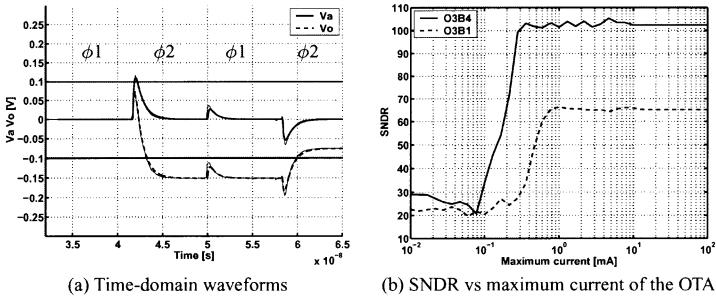


Figure 8.11: Influence of the SR on the performance of the $\Delta\Sigma$ converter. The left graph shows the waveforms of the input and output of the OTA. The model closely matches the waveforms from a full circuit simulation, indicated by the thin lines. The right graph compares the SR requirements for single-bit and multi-bit converters.

8.4.4 Slew-rate effects

Another important non-ideal effect in switched-capacitor integrators is the slewing of the OTA. The slewing performance of the integrator is dependent on many factors and it is very important to include all of them in order to obtain an accurate model.

First, the parasitic capacitances C_P and C_L , shown in Fig. 8.5, have a big influence on the slewing behavior of the integrator. Slewing is most likely to occur at the beginning of the integration phase. The voltage sampled on the sampling capacitance is then switched between the feedback signal and the input terminal of the OTA, causing a large voltage spike on node V_a which can drive the OTA into slewing. When C_P and C_L are not included in the model, the height of this spike equals $V_i - V_{fb}$ and slewing will occur very frequently. However, when C_P and C_L are included, the height of this initial voltage is significantly reduced since an immediate charge redistribution among the capacitors occurs at the beginning of the integration phase. The second requirement for the slew-rate model is the need to include the sampling operation of the next integrator. The waveforms of Fig. 8.11a illustrate this. At the beginning of ϕ_1 , C_{Sn} is connected to the output of the OTA and a voltage drop of the output is observed due to an immediate charge redistribution. Note that V_a shows the same drop, so the charge on the integration capacitance is not affected. Due to this peak, the OTA can also enter the slewing region during ϕ_1 . To ensure that the correct voltage is sampled on C_{Sn} , this peak should also settle. Therefore, the sampling operation of the next stage needs to be included in the model. Finally, the

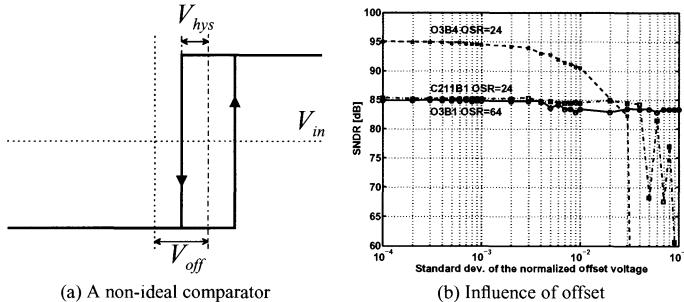


Figure 8.12: Influence of offset and hysteresis of the quantizer on the performance of $\Delta\Sigma$ converters. The simulations show a third-order $\Delta\Sigma$ converter with a one and four-bit quantizer and a single-bit cascaded 2-1-1 converter.

resistance of the switches smooths the initial peaks at the beginning of the clock-phases and thus also influence the slewing behavior.

Fig. 8.11a shows the transient waveforms of the input and output node of the integrator. The full horizontal lines indicate the limits for the slewing condition, given by $|V_a| > \frac{I_{\max}}{g_m}$. The thin lines are the waveforms from a full circuit simulation. This clearly shows that the model matches the circuit simulator very closely.

Fig. 8.11b shows simulation results for a third-order $\Delta\Sigma$ converter with one or four bits in the quantizer and an oversampling ratio of 32. A large difference can be observed between the single and multi-bit converters for the slew-rate simulations. In a multi-bit converter, the feedback signal tracks the input signal much closer and therefore the initial voltage drop at the input of the OTA will be much smaller. This results in more relaxed slew-rate specifications. Finally, it should be noted that the slew-rate specification depends a lot on the applied input frequency, especially for multi-bit $\Delta\Sigma$ converters. The reason is that a larger difference exists between the input and feedback signal as the signal frequency increases. This results in larger initial voltage steps at the input of the OTA and consequently slewing occurs more frequently. Therefore, larger slew-rate values are required to reduce the time spent in slewing and to ensure adequate settling in each clock phase.

8.4.5 Offset effects in the quantizer

The non-idealities in the quantizer are less important than the non-idealities of the feedback DAC due to the location in the $\Delta\Sigma$ converter. Any non-ideality of the DAC immediately appears unattenuated at the output of the $\Delta\Sigma$ converter. In contrast to that, the non-idealities of the quantizer are

suppressed by the gain of the preceding integrators. In fact, they are subject to the same noise-shaping action as the quantization noise. Therefore, they are generally less important and can be neglected in many cases. However, it will be shown that they can become a performance limiting factor in high-resolution $\Delta\Sigma$ converters [Gee 00b].

The quantizer in a $\Delta\Sigma$ converter runs at the same speed as the converter without any latency. Therefore, it is implemented as a flash AD converter consisting of several parallel comparators and a reference ladder to generate the required voltage taps. Fig. 8.12a shows the transfer function of a comparator including offset and hysteresis effects. In this chapter, only the influence of offset is discussed. The offset is considered a random variable, which depends on the matching performance of the technology and the sizes and topology of the comparator. Monte-Carlo simulations of the behavioral model are performed for various $\Delta\Sigma$ converters. The worst-case results are shown in Fig. 8.12b.

Fig. 8.12b shows that the third-order single-bit converter is very insensitive to these non-idealities. For a reference voltage of 1V, the standard deviation of the offset needs to be smaller than 100mV to have less than 3dB degradation. For the cascaded converter, this specification is 40mV. These simulations show that single-bit $\Delta\Sigma$ converters are very insensitive to non-idealities in the quantizer. The offset specification requires some care during the design, but it imposes no real problems.

When the same simulations are performed for a third-order four-bit $\Delta\Sigma$ converter, the situation is quite different. The standard deviation of the offset voltage should be smaller than 6mV. Due to the offset specification, fairly large input transistors are required which results in a larger input capacitance of the comparator. On top of that, the multi-bit quantizer has a number of comparators in parallel which increases the input capacitance even further. This total input capacitance increases the effective load capacitance of the last integrator. Together with the settling requirements, this can lead to an increased power consumption of the last integrator [Gee 00b].

8.5 Design example

In this section, the design of a single-loop multi-bit $\Delta\Sigma$ converter is presented [Gee 00b, Gee 00a]. The system diagram is shown in Fig. 8.13. A third-order four-bit topology is selected which can achieve a resolution of 16 bits for an oversampling ratio of only 24. Data Weighted Averaging is used to limit the accuracy requirement of the DA converter in the feedback loop. The output of this DWA block has to be distributed over the entire

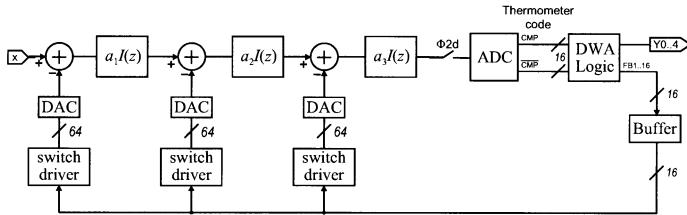
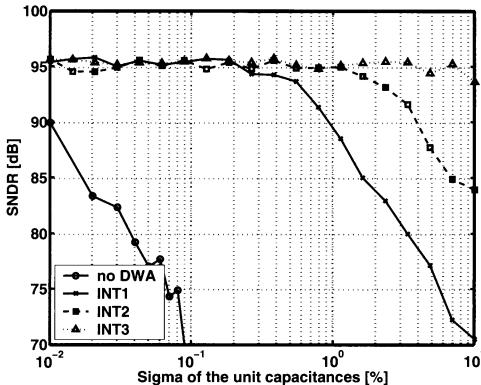
Figure 8.13: System diagram of the multi-bit $\Delta\Sigma$ converter.

Figure 8.14: Monte-Carlo simulations to determine the required matching of the unit capacitances of each stage.

chip to the local switch drivers. Therefore, a buffer is inserted to deal with the large gate and wiring capacitance. The local switch drivers generate the local control signals for the switches of the DACs, which are implemented together with the integrators as shown in Fig. 8.4b.

A behavioral model of the $\Delta\Sigma$ converter with DWA algorithm and a non-ideal DAC is used for Monte-Carlo simulations. Fig. 8.14 shows the worst case SNDR of the $\Delta\Sigma$ converter versus the required matching of the unit capacitances of the different integrators. The first curve shows the performance without the DWA algorithm for the first integrator, while the other curves include the DWA algorithm in the model. These simulations clearly show the reduced linearity requirement of the DAC due to the DWA algorithm. For the first integrator, a sigma of 0.2% is required, while the second and third integrator only need a sigma of 1% and 10%, respectively. This clearly shows that the non-idealities of the 2nd and 3rd integrator are suppressed by the gain of the preceding integrators.

Table 8.3: Specifications and designed values for the different building blocks of the converter.

	Specification	Designed Value
OTA gain	> 80 dB	80dB
Dominant closed-loop pole	$> 1.15 \cdot 2.5 \cdot f_s = 172\text{MHz}$	220MHz
Switch resistance	$< \frac{1}{1.15 \cdot 2.5 \cdot f_s \cdot 2\pi \cdot C_S} = 288\Omega$	220Ω
Slew-Rate	$> 60 \text{ V}/\mu\text{s}$	180V/ μ s

The size of the unit capacitance of the first integrator is determined by kT/C noise requirements or by the matching requirements for the feedback DAC. Using (8.3), it can be calculated that a 3.2pF sampling capacitance results in SNR_{kT/C} of 99dB since the reference voltage of the converter is 1V. This results in a 200fF unit capacitance. The sigma of the matching of these unit capacitances is better than 0.2%. This means that the sampling capacitance is still determined by the kT/C noise requirements and not by the matching requirements of the DAC thanks to the use of the DWA algorithm.

The specifications for the different building-blocks are derived from behavioral simulations, as discussed in the previous sections. Table 8.3 shows some of the most important specifications. Note that a 15% margin is added to the specification of the dominant closed-loop pole and the resistance of the switches to take the lost time due to the non-overlapping clocks into account. A folded-cascode OTA with gain-boosting stages is used to combine a large gain with an excellent frequency performance.

The converter is implemented in a 0.65 μ m CMOS process, operating from a 5V supply. The location of the different building blocks is indicated on the micro-photograph of the chip, shown in Fig. 8.15. Special care has been taken to provide identical surroundings for the unit capacitances of the DAC and to shield the most sensitive nets such as the reference voltages. The total area, including bonding pads and decoupling capacitances, is 5.3mm².

Fig. 8.16a shows the output spectrum, while Fig. 8.16b shows the measured SNR and SNDR of respectively 95 and 89dB for a clock frequency of 60MHz and an oversampling ratio of 24. A dynamic range of 97dB is achieved in a 1.25MHz signal bandwidth. The power consumption is 295mW, of which 152mW is consumed in the analog part. The digital power consumption is mainly due to the clock buffer, which generates the non-overlapping clocks for the switch-capacitor circuits from one external clock signal. The most important measurement results are summarized in Table 8.4.

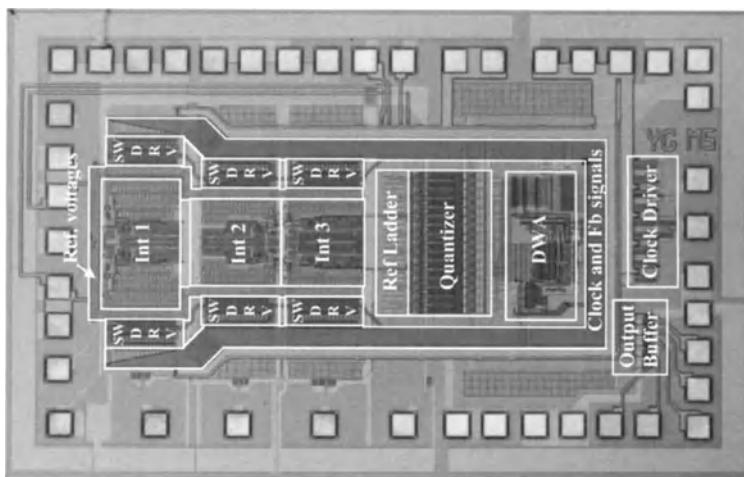


Figure 8.15: Micro-photograph of the converter.

Table 8.4: Summary of the measurement results.

Dynamic Range	97 dB
SNR	95 dB
SNDR	89 dB
STHD	97 dB
Output Rate	2.5 MS/s
Clock Frequency	60 MHz
Oversampling Ratio	24
Power Consumption:	
Analog	152 mW
Digital	143 mW
Total	295 mW
Supply	5V
Technology	0.65 μ m CMOS TMDP
Die size	1.83 x 2.92 mm ²

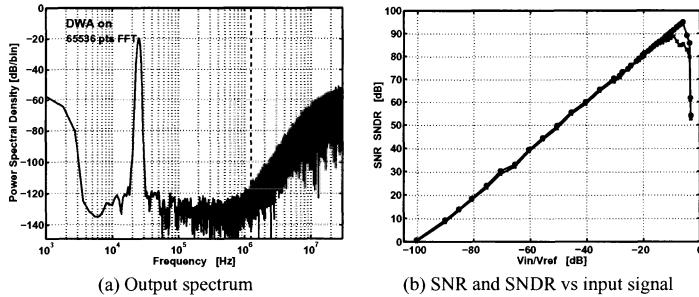


Figure 8.16: Measurement results of the converter.

8.6 Conclusion

$\Delta\Sigma$ converters are widely applied to achieve high-performance AD conversion. A brief review of architectures was discussed to show that either multi-bit and/or cascaded converters are most suited to combine a high-speed with a high-resolution. The linearity requirements of the DAC in the feedback loop of multi-bit converters can be significantly reduced by employing Dynamic Element Matching techniques, without sacrificing any of the advantages of multi-bit converters.

In order to obtain a well designed implementation of these converters, it is important to know which are the performance limiting circuit non-idealities. In that respect, the (non-linear) resistance of the switches, the gain, settling and slewing of the OTA and the offset of the quantizer have been discussed. A 16-bit 2.5MS/s $\Delta\Sigma$ converter which uses a single-loop multi-bit third-order architecture with DEM has been discussed as a design example.

References

- [Abo 99] A. M. Abo and P. R. Gray, “A 1.5V, 10-bit, 14.3-MS/s CMOS Pipeline Analog-to-Digital Converter”, *IEEE Journal of Solid-State Circuits*, vol. SC-34, no. 5, pp. 599–606, May 1999.
- [Ada 86] R. W. Adams, “Design and Implementation of an Audio 18-Bit Analog-to-Digital Converter using Oversampling Techniques”, *Journal Audio Eng. Soc.*, vol. SC-34, pp. 153–166, Mar. 1986.
- [Ada 96] R. W. Adams and R. Schreier, “Stability Theory for Delta-Sigma Modulators”, in *Delta-Sigma Data Converters: theory, design, and simulation*, S. Norsworthy, R. Schreier, and G. Temes, Eds. 1996, IEEE Press.

- [Bai 94] R. T. Baird and T. S. Fiez, “Stability Analysis of High-Order Delta-Sigma Modulation for ADC’s”, *IEEE Transactions on Circuits and Systems II*, vol. 41, no. 1, pp. 59–62, Jan. 1994.
- [Bai 95] R. T. Baird and T. S. Fiez, “Linearity Enhancement of Multibit $\Sigma\Delta$ A/D and D/A Converters Using Data Weighted Averaging”, *IEEE Transactions on Circuits and Systems II*, vol. 42, no. 12, pp. 753–762, Dec. 1995.
- [Bai 96] R. T. Baird and T. S. Fiez, “A Low Oversampling Ratio 14-b 500-kHz Delta-Sigma ADC with a Self-Calibrated Multibit DAC”, *IEEE Journal of Solid-State Circuits*, vol. 31, no. 3, pp. 312–320, Mar. 1996.
- [Bra 91] B. P. Brandt and B. A. Wooley, “A 50MHz multi-bit Sigma-Delta Modulator for 12-b 2-MHz A/D Conversion”, *IEEE Journal of Solid-State Circuits*, vol. 26, no. 12, pp. 1746–1756, Dec. 1991.
- [Bro 97] T. L. Brooks, D. H. Robertson, D. F. Kelly, A. D. Muro, and S. W. Harston, “A Cascade Sigma-Delta Pipeline A/D Converter with 1.25 MHz Signal Bandwidth and 89 dB SNR”, *IEEE Journal of Solid-State Circuits*, vol. 32, no. 12, pp. 1896–1906, Dec. 1997.
- [Bul 00] K. Bult, “Analog Design in Deep Sub-Micron CMOS”, in *Proceedings European Solid-State Circuits Conference*, Sept. 2000, pp. 11–17.
- [Can 96] J. C. Candy, “An Overview of Basic Concepts”, in *Delta-Sigma Data Converters: theory, design, and simulation*, S. Norsworthy, R. Schreier, and G. Temes, Eds. 1996, IEEE Press.
- [Car 96] R. L. Carley, R. Schreier, and G. C. Temes, “Delta-Sigma ADCs with Multibit Internal Converters”, in *Delta-Sigma Data Converters: theory, design, and simulation*, S. Norsworthy, R. Schreier, and G. Temes, Eds. 1996, IEEE Press.
- [Cat 89] T. Cataltepe, A. R. Kramer, L. E. Larsen, G. C. Temes, and R. H. Walden, “Digitally corrected multi-bit $\Sigma\Delta$ data converters”, in *Proceedings IEEE International Symposium on Circuits and Systems*, 1989, pp. 647–650.
- [Cho 95] T. B. Cho and P. R. Gray, “A 10b, 20Msample/s, 35mW Pipeline A/D Converter”, *IEEE Journal of Solid-State Circuits*, vol. 30, pp. 166–172, Mar. 1995.
- [Des 01] M. Dessouky and A. Kaiser, “Very Low-Voltage Digital-Audio $\Delta\Sigma$ Modulator with 88 dB Dynamic Range Using Local Switch Bootstrapping”, *IEEE Journal of Solid-State Circuits*, vol. 36, no. 3, pp. 349–355, Mar. 2001.
- [Fat 93] J. W. Fattaruso, S. Kiriaki, M. de Wit, and G. Warwar, “Self-Calibration Techniques for a Second-Order Multibit Sigma-Delta Modulator”, *IEEE Journal of Solid-State Circuits*, vol. SC-28, no. 12, pp. 1216–1223, Dec. 1993.

- [Fuj 00] I. Fujimori, L. Longo, A. Hairapetian, K. Seiyama, S. Kosic, J. Cao, and S.-L. Chan, “A 90-dB SNR 2.5-MHz Output-Rate ADC Using Cascaded Multibit Delta-Sigma Modulation at 8x Oversampling Ratio”, *IEEE Journal of Solid-State Circuits*, vol. 35, no. 12, pp. 1820–1828, Dec. 2000.
- [Gee 99] Y. Geerts, A. Marques, M. Steyaert, and W. Sansen, “A 3.3 V 15-bit Delta-Sigma ADC with a Signal Bandwidth of 1.1MHz for ADSL-applications”, *IEEE Journal of Solid-State Circuits*, vol. 34, no. 7, pp. 927–936, July 1999.
- [Gee 00a] Y. Geerts, M. Steyaert, and W. Sansen, “A 2.5MSample/s Multi-Bit $\Delta\Sigma$ CMOS ADC with 95dB SNR”, in *Proceeding International Solid-State Circuits Conference*, San Francisco, Feb. 2000, pp. 336–337.
- [Gee 00b] Y. Geerts, M. Steyaert, and W. Sansen, “A High-Performance Multi-Bit CMOS $\Delta\Sigma$ Converter”, *IEEE Journal of Solid-State Circuits*, vol. 35, no. 12, pp. 1829–1840, Dec. 2000.
- [Gee 02] Y. Geerts, M. Steyaert, and W. Sansen, *Design of Multi-Bit Delta-Sigma A/D Converters*, Kluwer Academic Publishers, 2002.
- [Gro 89] D. W. J. Groeneveld, H. J. Schouwenaars, H. A. H. Termeer, and C. A. A. Bastiaansen, “A Self-Calibration Technique for Monolithic High-Resolution D/A Converters”, *IEEE Journal of Solid-State Circuits*, vol. SC-24, no. 6, pp. 1517–1522, Dec. 1989.
- [Hai 91] A. Hairapetian, G. C. Temes, and Z. X. Zhang, “Multibit sigma-delta modulator with reduced sensitivity to DAC nonlinearity”, *Electronics Letters*, vol. 27, pp. 990–991, May 1991.
- [Lak 94] K. R. Laker and W. M. Sansen, *Design of Analog Integrated Circuits and Systems*, McGraw-Hill, New York, 1994.
- [Mar 98a] A. Marques, V. Peluso, M. Steyaert, and W. Sansen, “A 15-b Resolution 2MHz Nyquist Rate $\Delta\Sigma$ ADC in a 1- μm CMOS Technology”, *IEEE Journal of Solid-State Circuits*, vol. 33, no. 7, pp. 1065–1075, July 1998.
- [Mar 98b] A. Marques, V. Peluso, M. Steyaert, and W. Sansen, “Optimal Parameters for $\Delta\Sigma$ modulator topologies”, *IEEE Transactions on Circuits and Systems*, vol. 45, no. 9, pp. 1232–1241, Sept. 1998.
- [Med 99a] F. Medeiro, B. Perez-Verdu, and A. Rodriguez-Vazquez, “A 13-bit, 2.2 MS/s, 55-mW Multibit Cascade $\Sigma\Delta$ Modulator in CMOS 0.7- μm Single-Poly Technology”, *IEEE Journal of Solid-State Circuits*, vol. 34, no. 6, pp. 748–760, June 1999.
- [Med 99b] F. Medeiro, L. Perez-Verdu, and A. Rodriguez-Vazquez, *Top-down design of high-performance Sigma-Delta Modulators*, Kluwer Academic Publishers, 1999.

- [Nor 96] S. Norsworthy, R. Schreier, and G. Temes, Eds., *Delta-Sigma Data Converters: theory, design, and simulation*, IEEE Press, 1996.
- [OpT 93] F. Op'T Eynde and W. Sansen, *Analog interfaces for digital signal processing systems*, Kluwer Academic Publishers, 1993.
- [Pan 00] H. Pan, M. Segami, M. Choi, J. Cao, and A. A. Abidi, “A 3.3V, 12b, 50MS/s A/D Converter in $0.6\mu\text{m}$ CMOS with over 80dB SFDR”, *IEEE Journal of Solid-State Circuits*, vol. SC-35, no. 12, pp. 1769–1780, Dec. 2000.
- [Pel 97] V. Peluso, M. Steyaert, and W. Sansen, “A 1.5V - $100\mu\text{W}$ delta-sigma Modulator with 12-b Dynamic Range Using the Switched-Opamp Technique”, *IEEE Journal of Solid-State Circuits*, vol. 32, no. 7, pp. 943–952, July 1997.
- [Rad 00] R. E. Radke, A. Eshraghi, and T. S. Fiez, “A 14-bit Current-Mode $\Sigma\Delta$ DAC Based Upon Rotated Data Weighted Averaging”, *IEEE Journal of Solid-State Circuits*, vol. 35, no. 8, pp. 1074–1084, Aug. 2000.
- [Raz 95] B. Razavi, *Principles of Data Conversion System Design*, IEEE Press, 1995.
- [Rib 91] D. B. Ribner, “A comparison of modulator networks for high-order oversampled $\Sigma\Delta$ analog-to-digital converters”, *IEEE Transactions on Circuits and Systems*, vol. 38, no. 2, pp. 145–159, Feb. 1991.
- [Sar 93] M. Sarhang-Nejad and G. Temes, “A high resolution multibit $\Sigma\Delta$ ADC with digital correction and relaxed amplifier requirements”, *IEEE Journal of Solid-State Circuits*, vol. 28, no. 6, pp. 648–660, June 1993.
- [Tem 80] G. C. Temes, “Finite Amplifier Gain and Bandwidth Effects in Switched-Capacitor Filters”, *IEEE Journal of Solid-State Circuits*, vol. SC-15, no. 3, pp. 358–361, June 1980.
- [Yin 93] G. Yin, F. Stubbe, and W. Sansen, “A 16-bit 320kHz CMOS A/D Converter Using Two-Stage Third-Order $\Sigma\Delta$ Noise-Shaping”, *IEEE Journal of Solid-State Circuits*, vol. 28, pp. 640–647, June 1993.

Chapter 9

High-order Cascade Multi-bit $\Sigma\Delta$ Modulators

*F. Medeiro, R. del Río, J.M. de la Rosa, B. Pérez-Verdú
and A. Rodríguez-Vázquez*

9.1 INTRODUCTION ^{†1}

Motivated by the commercial success of the wireline communication products, mixed-signal designers are being pushed to integrate A/D and D/A interfaces featuring 12- to 16-bit effective resolution for signal bandwidths well in excess of 1MHz [1]. These specifications must be achieved in a low-voltage scenario, making use of poor performance (and often badly characterized) devices, which decreases the “analog speed” of deep-submicron CMOS processes.

In this context, oversampled Sigma-Delta modulation ($\Sigma\Delta$) [2] is usually preferred to other A/D conversion techniques for its low-complexity analog circuitry and robustness. Since they trade time-domain redundancy and digital processing for analog performance of the building block, these converters are relatively easily integrated in modern deep-submicron CMOS processes, more suited to implement fast (either analog or digital circuit and devices) than precise analog functions. However, the latter is absolutely true only if the oversampling ratio (M) –i.e., the ratio between the actual sampling frequency and the Nyquist frequency of the signal being converted – is high, which obviously cannot be the case in high-speed communication. As a matter of example, if the signal bandwidth is 1.1MHz (like in ADSL modems [1]), so that the Nyquist frequency is 2.2MHz, an oversampling ratio of 256, commonly found in audioband processing, will require 563.2MHz sampling frequency! So, a moderate amount of oversampling is mandatory. In fact, in spite of the increasing potential speed of the new CMOS processes, the trend is to decrease M up to 16 and below, because of the concurrent increase of the converter bandwidth [3]-[10].

1. **Acknowledgement:** This work has been funded by following institutions: EU (IST Project 29261/MIXMODEST and IST Project 2001-34283/TAMES-2), and the Spanish MCyT and the ERDF (Project TIC2001-0929/ADAVERE).

The oversampling ratio largely influences the performance of a $\Sigma\Delta$ M, but there are two other important design parameters, namely: the modulator order (L) and the resolution of the internal quantizer (B). For the sake of clarity, these parameters have been associated to their respective blocks in the simplified $\Sigma\Delta$ converter schematic in Fig. 9.1. By using additive-error, linearized models [2], simple z-domain algebra shows that, under ideal operating conditions, the dynamic range (DR) of a $\Sigma\Delta$ M is given by:

$$DR_{dB} = 10 \log \left[\frac{3}{2} (2^B - 1)^2 (2L + 1) \cdot M^{2L+1} \right] \quad ENOB = \frac{DR_{dB} - 1,76}{6,02} \quad (9.1)$$

Out of the three main design parameters, increasing the oversampling ratio has been the most popular resort for augmenting the modulator resolution, the reason being twofold: (a) On the one hand M has a clear, beneficial impact on resolution^{†2}, (b) on the other, while changing M basically alters the dynamic requirements of the building blocks^{†3}, augmenting L and/or B also raises issues at the modulator level, jeopardizing loop stability and/or degrading robustness [2].

Unfortunately, as shown previously, the maneuverability of the oversampling ratio is certainly restricted in a broadband telecom converter, so that in order to cope with oversampling ratios below 32, high-order filtering and/or multi-bit quantization must be used. In doing so, the original robustness of the highly-over-sampled low-order single-bit $\Sigma\Delta$ conversion is weakened, which often obligates to

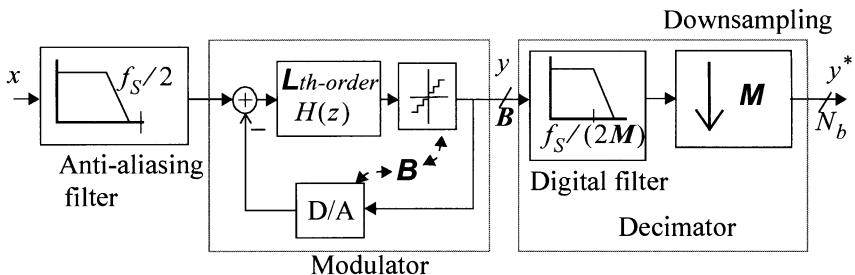


Figure 9.1: Block diagram of a $\Sigma\Delta$ A/D converter.

-
2. According to eq. (9.1), increasing M generates an increase in resolution of $L+0.5$ bit/octave. For example, for a fourth-order modulator ($L = 4$) the effective resolution is enhanced in 4.5bit each time M is doubled. For the same example with $M = 16$, resolution would increase in approximately 2.5bit per each unitary increase in L , whereas roughly an extra bit is obtained each time B is increased.
 3. Obviously, it also affects the operation of the decimation filter, but this block is out of the scope of the chapter. The interested reader can consult Chapter 12 in this book.

resort to correction/calibration mechanisms and, definitely, to a more careful mixed-signal design. In short, augmenting L or B generates two main problems:

- Unlike first- and second-order loops, high-order loops are not unconditionally stable.
- The non-linearity of the multi-bit DAC in the feedback path in Fig. 9.1 does manifest as non-shaped and, hence, non-attenuated error. Note that it could never happen in a single-bit case because a 2-level (1-bit) DAC is intrinsically linear^{†4}.

Both problems have been alleviated. High-order $\Sigma\Delta$ Ms can be stabilized through several techniques. Conditionally stable modulators have been designed by properly choosing the scaling factors [12]; or by using multi-path feedforward structures to reduce the out-of-band value of the quantization noise transfer function [13]. Also, instabilities can be circumvented by resetting some integrator outputs if divergencies are detected [14]. However, mechanisms adopted for stabilization generally deteriorate the dynamic range with respect to that predicted by eq. (9.1). Reductions of around 25dB in DR (approximately 4bits) and even more have been reported [2]. Furthermore, stabilization often requires very small scaling coefficients, which means very large capacitors in SC implementations - unhelpful for power saving. Finally, the complexity of the modulator architecture increases considerably, which may result in non-optimum designs.

Regarding the multi-bit DAC non-linearity, several correction/calibration methods have been employed working either in the digital domain [15][16] or in the analog domain [17]-[19]. Among the latter, dynamic element matching techniques [17][19][20] are interesting because they enable on-line correction and have been successfully applied at high frequencies [8][9]. However, since DACs cannot be efficiently linearized within an arbitrarily large resolution, the use of low-order multi-bit modulation may not be enough to obtain a given DR .

A direct solution to this problem is to increase both the modulator order and the internal quantization resolution, giving rise to moderate-order (3 - 5), multi-bit architectures. In fact, the use of multi-bit quantization in single-loop high-order modulators inherently improves their stability properties [2], so that these are good candidates to obtain high-resolution, high-frequency operation, provided that the non-linearity problem is solved.

With the same objective, the combination of high-order cascade (MASH) architectures [21] with multi-bit quantization has been proposed [22]-[25]. These modulators gather the unconditional stability of cascade modulators (provided that only second- and/or first-order stages are used) and the advantages of multi-

4. It has been shown [11] that only linearity matters. Other DAC errors, such as offset and gain, have an impact only for extreme values, uncommon in a practical design.

bit quantization with relaxed requirements for the linearity of the latter. The feasibility and efficiency of this approach, because it needs no correction/calibration mechanism, has been proved in analog technologies [3][4][22]. Further investigation into the potentialities of these architectures have led us to propose an easily expandible, modular family of high-order cascades. As explained in Section 9.2, thanks to a proper selection of coefficients, these modulators preserve their low systematic loss and high overloading point regardless of the order, with relaxed output swing requirement.

Circuit imperfections affecting cascade modulators can hide some of the benefits of multi-bit quantization. Particularly, finite amplifier DC-gain and integrator weight mismatch cause low-order quantization error leakage that may degrade the *DR*. Paradoxically, some design problems worsen with the technology advances: Firstly, decreasing the supply voltage constrains more and more the dynamic range and invalidates some popular circuit techniques. Secondly, simultaneously achieving high gain and speed in amplifiers (the most important block in a SC $\Sigma\Delta M$) is getting harder due to the transconductance and output conductance degradation in deep-submicron MOS devices. Resorting to more sophisticated circuit techniques can help to circumvent these problems, but often at the price of an increased power consumption, which contradicts the initial assumption: the smaller the technology, the lower the power dissipation. Overall, in the short term there will be an inflection of the curve *power dissipation vs. technology scaling*, as we show in Section 9.3, aimed at getting the $\Sigma\Delta M$ architectures covered in this chapter in a technological perspective.

Despite these problems, we conclude that cascade multi-bit modulators can be efficiently implemented in deep-submicron CMOS processes, exhibiting similar or even superior performance to other $\Sigma\Delta M$ architectures. Moreover, this can be achieved by employing simple, non-calibrated analog circuitry in keeping with the original $\Sigma\Delta$ modulation philosophy. The key to success is a careful mixed-signal design assisted by precise models of the circuit behavior. To demonstrate this, Section 9.4 is devoted to describe a practical design example.

9.2 EXPANDIBLE, MODULAR CASCADE MULTI-BIT MODULATORS

9.2.1 Cascade multi-bit $\Sigma\Delta$ modulators

Fig. 9.2 shows the block diagram of a generic cascade multi-bit $\Sigma\Delta$ modulator. A cascade (MASH) modulator [21] consists of a number N of low-order modulators or stages connected in such a way that each stage modulates a signal proportional to the quantization error generated in the previous one. Once in the digital domain, the quantization errors of all stages, but the last one, are cancelled by

simple algebraic operations and delays. The last-stage quantization error appears at the modulator output shaped by a function of order equal to the summation of the orders of all stages. Since the modulator input is unaffected by this procedure, the performance of a cascade modulator is equivalent to that of an ideal high-order loop. Furthermore, provided that only second- and/or first-order modulators are used, the cascade can be designed to be unconditionally stable.

A robust multi-bit $\Sigma\Delta M$ is obtained by increasing the number of quantization levels in some of the quantizers of a cascade architecture [11][22]-[25]. In the case of the architectures covered in this chapter, multi-bit quantization is included only in the last stage, while the remaining are single-bit. Assuming perfect cancellation of the low-order quantization error, it can be shown that only the signal $X(z)$ and the last-stage quantization error $E_N(z)$ remain at the z -domain modulator output, yielding:

$$Y(z) = z^{-L} \cdot X(z) + d \cdot (1 - z^{-1})^L \cdot E_N(z) \quad L = L_1 + L_2 + \dots + L_N \quad (9.2)$$

Note that since $E_N(z)$ is generated in a B -bit quantizer, the modulator response equals that of an ideal L th-order B -bit $\Sigma\Delta M$, except for a scalar d , equal to the inverse of the product of some integrator weights in the cascade. In order not to prematurely overload the modulator, signals must be properly scaled down as they are transmitted from one stage to the next. This results in a value for d larger than unity, which means amplification of the last-stage quantization error, thus decreasing the effective resolution in $\log_2(d)$ bits. Due to its large impact, minimizing the value of d must take priority during the architecture selection task.

Moreover, using multi-bit quantization only in the last stage relaxes the linearity requirement of the multi-bit DAC. In this case, the non-linearity error is not

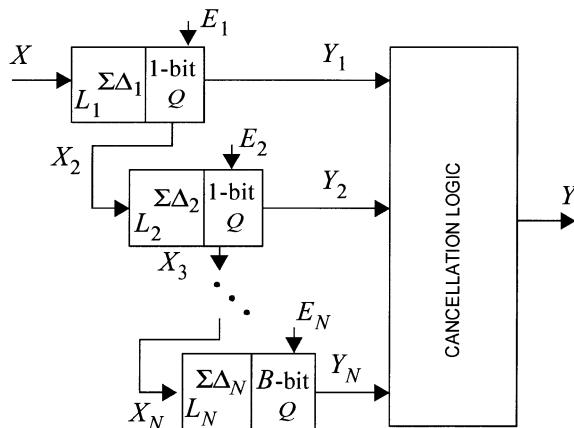


Figure 9.2: Generic cascade single-bit/multi-bit $\Sigma\Delta$ modulator.

injected at the modulator input. This allows us to high-pass filter the DAC errors, so that most part of its power is pushed out of the signal band. Simple analysis shows,

$$Y(z) = z^{-L} \cdot X(z) + d \cdot (1 - z^{-1})^L \cdot E_N(z) + d \cdot (1 - z^{-1})^{(L - L_N)} \cdot E_{INL}(z) \quad (9.3)$$

where E_{INL} is the z -transform of the integral non-linearity error in the last-stage DAC. Such an error presents a shaping function of order $L - L_N$ (i.e., the overall modulator order minus the order of the last stage). It means that such non-ideality can be tolerated to some extent without correction/calibration.

Despite their theoretical good characteristics, the successful design of cascade multi-bit $\Sigma\Delta$ modulators requires some considerations related to both architecture and circuitry, which are analyzed next.

9.2.2 Architecture optimization

There are three degrees of freedom in the generic architecture of Fig. 9.2, namely: (a) total number of stages, N , (b) order of each stage, $L_1 \dots L_N$, and (c) number of bits of the last-stage quantization, B . Recall that selecting the number of stages and their respective order is equivalent to setting the overall modulator order, $L = L_1 + L_2 + \dots + L_N$. As explained in the introduction of this chapter, there is another basic parameter at the architectural level: the oversampling ratio (M). In this section we will focus on the selection of the number of stages and their orders, because they are directly related to the architecture optimization. The selection of M and B will be considered further on.

In order to avoid instability issues, we limit to 1 or 2 the possible orders. It has been shown that in order to gain insensitivity to circuit imperfections, the first stage must be second-order [11][22]; so we will always consider $L_1 = 2$. The remaining stages can be either first- or second-order. Thus the immediate questions are: *How many stages?* and *First-, second-order, or a combination of both?*

In order to elucidate these questions, recall that signals transmitted along the cascade must be properly scaled to prevent premature overload of the stages. The larger the number of stages, the more acute the overloading problem becomes. Moreover, it is known that, unlike first-order modulators, second-order ones overload before reaching the full-scale input. This means that, if the second and successive stages are second-order, more attenuation will be needed in the signal being transmitted from stage to stage. As mentioned before, digitally compensating such scale factors will result in a higher value of the coefficient d in eqs. (9.2) and (9.3) and, hence, higher systematic loss of resolution. We can conclude that the minimum value of the digital coefficient d is larger the larger the number of stages and/or when second-order stages are used. This qualitative reasoning,

which will be quantitatively supported further on, lead us to adopt the architecture shown in Fig.9.3. It is a L th-order modulator formed by a second-order stage followed by $L - 2$ identical first-order stages. The values of their integrator weights are,

$$\begin{aligned} g_{1a} &= g_{1a}' = 0,25 & g_{1b} &= 1, g_{1b}' = 0,5 \\ g_k &= 1, g_k' = 0,5, g_k'' = 0,5 & k &= 2, \dots, L-1 \end{aligned} \quad (9.4)$$

and the digital blocks are:

$$H_k(z) = (1 - z^{-1})^k, \quad k = 2 \dots L-1 \quad (9.5)$$

With coefficients in eq. (9.4) the value of the scalar d in eqs. (9.2) and (9.3) turns out to be 2, which means a systematic loss of resolution of 1bit. This loss in one of the smallest possible and considerably smaller than that of other high-order cascades, as we show at the end of this section. The set of coefficients above also have the following interesting properties: (a) The total output swing (*OS*) required in all integrators equals the quantizer full-scale. Such an appealing feature for low-voltage implementation is illustrated in Fig.9.4. (b) The largest weight of each 3-weight integrator can be obtained as the summation of the others, so that no 3-branch SC integrator is required. This also minimizes the number of unitary

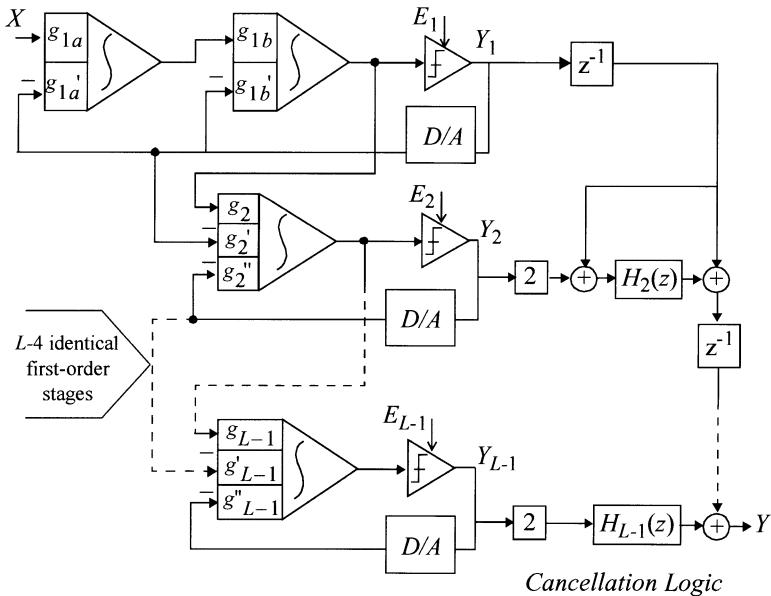


Figure 9.3: L th-order $\Sigma\Delta$ modulator using a $2-1^{(L-2)}$ cascade.

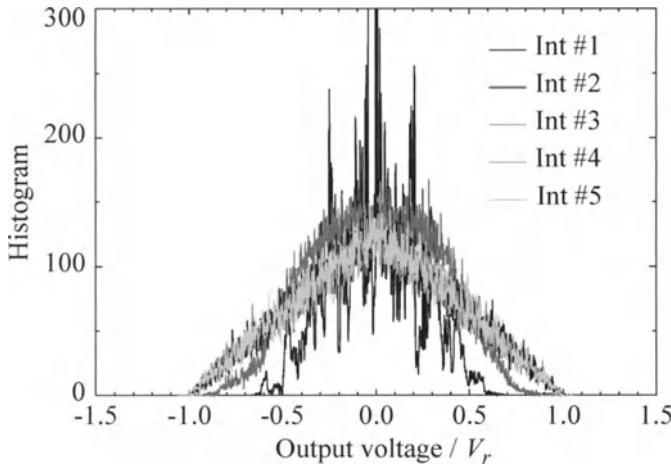


Figure 9.4: Histogram of the integrator output voltage relative to the reference voltage for a fifth-order cascade ($L = 5$), with coefficients in eq. (9.4).

capacitors. (c) All first-order stages use the same analog and digital coefficients, and they can be electrically identical as well. This considerably simplifies the electrical and physical implementation of the converter.

However, probably the most appealing feature of the architecture in Fig. 9.3 with the set of coefficients in eq. (9.4) is that it can be easily set to any order just by changing the number of identical first-order stages. In this process a correct operation is maintained with constant overloading point. This is another good characteristic of the set of coefficients proposed, since normally the trend to premature overloading increases with the number of stages. This is illustrated in Fig. 9.5 where the ideal signal-to-(noise+distortion) ratio (*SNDR*) curves for $L = 4$, 5, 6, and 7 are plotted. Note that, although the number stages is increased from curve to curve, the overloading point does not change.

In order to reinforce our architecture and coefficient selection, Fig. 9.6 shows the *SNDR* vs. input level for two versions of the more usual 6th-order 2-2-2 cascade [4][26], and the one in this chapter with $L = 5$ and 6. Note that even with $L = 5$, the *SNDR* is improved with respect to the 6th-order 2-2-2 cascades. The reason is that, due to the use of second-order modulators in the second and third stages of the 2-2-2 architecture, more attenuation is needed for the signal transmitted between stages and a systematic loss of resolution between 4 and 6 bits is common. The following section shows that the superiority of the architecture here (under ideal conditions) is not jeopardized by the presence of circuit imperfections.

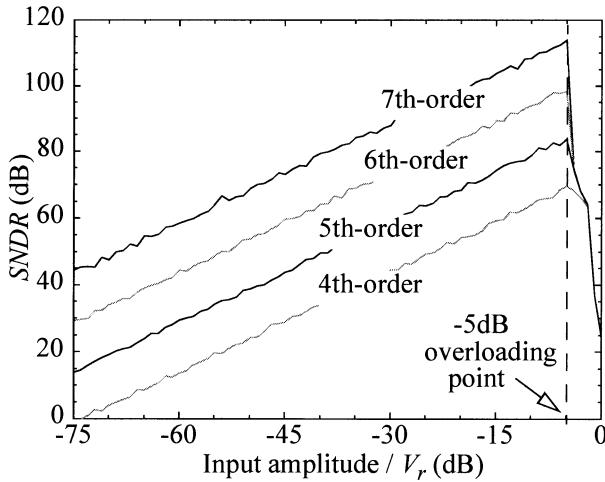


Figure 9.5: $SNDR$ for several versions of the expandible architecture. In all cases the oversampling ratio is 16.

Let's return to the other main parameters of the architecture, M and B . Remember that the main objective of using a high-order $\Sigma\Delta M$ is to reduce the oversampling ratio required to obtain a given resolution, which, according to eq. (9.1), can also be accomplished by augmenting B . For instance, Fig. 9.7 shows the minimum oversampling ratio needed to achieve 14bit (or 86-dB DR) as a function of the

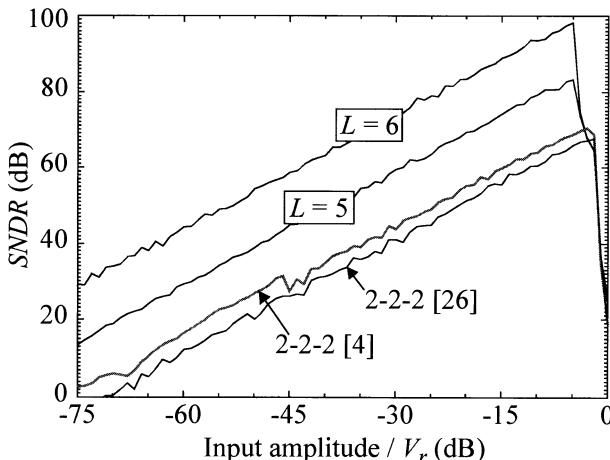


Figure 9.6: Ideal $SNDR$ for two versions of the 2-2-2 cascade and the architecture here with $L = 5$ and 6. In all cases the oversampling ratio is 16.

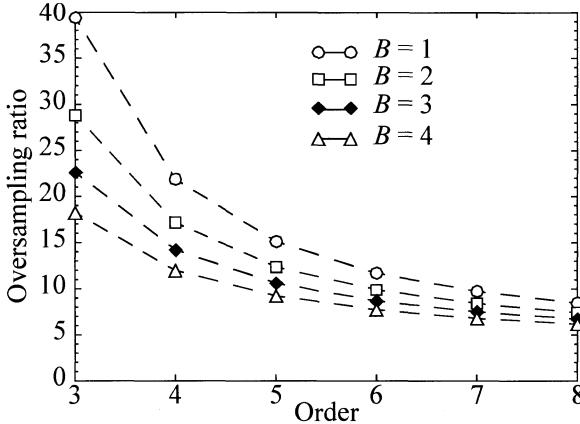


Figure 9.7: Oversampling ratio required to obtain 14bit with a 2-1-...-1 modulator of order ranging from 3 to 8. Curves for B larger than one are also shown.

order of the $2-1^k$ topology^{†5}. Note that, for a given order, further reduction of the oversampling can be achieved with $B > 1$. Thus, if we consider, for example, a fourth-order $2-1^2$ cascade, the required resolution can be achieved with $\times 16$ oversampling using 3bits in the last-stage quantizer, whereas its single-bit counterpart would require $\times 22$ oversampling.

9.3 TECHNOLOGY-DRIVEN SCENARIO FOR CASCADE MULTI-BIT SIGMA-DELTA MODULATORS

9.3.1 Non-ideal operation

Switched-capacitor implementations of cascade modulators suffer from certain non-ideal behaviors more than their single-loop counterparts, namely: finite (and non-linear) amplifier DC-gain and capacitor mismatch.

So far, we have shown that for a $2-1^kmb$ $\Sigma\Delta M$ the output error terms are the last-stage quantizer error with a shaping function of order L , plus the DAC non-linearity error with a shaping function of order $L-1$. This result is obtained from eq. (9.3) making $L_N = 1$. From that equation, the following expression for the approximate in-band error power can be derived [11]:

$$P_Q \equiv d^2 \cdot \sigma_{mQ}^2 \frac{\pi^{(2L)}}{(2L+1) \cdot M^{(2L+1)}} + d^2 \cdot \sigma_{INL}^2 \frac{\pi^{(2L-2)}}{(2L-1) \cdot M^{(2L-1)}} \quad (9.6)$$

5. $2-1^k$ stands for the modulator formed by cascading a second-order stage and k first-order stages. We will use $2-1^kmb$ for denoting its multi-bit version (only in the last stage).

where $\sigma_{mQ}^2 = [2V_r/(2^B - 1)]^2/12$ is the power of the last-stage quantization error ($2V_r$ stands for the last-stage quantizer full-scale, FS) and σ_{INL}^2 represents the power of the error generated by the DAC, which can be estimated as $\sigma_{INL}^2 \cong (1/2)(2V_r)^2(INL/100)^2$, with INL being the DAC integral non-linearity expressed in %FS.

The above expression and its origin, eq. (9.3), have been obtained assuming ideal operation of the integrator block, which is represented by its finite-difference equation,

$$v_{o,n} = \sum_{i=1}^{n_b} g_i v_{i,n-1} + v_{o,n-1} \quad (9.7)$$

valid for an integrator with n_b input branches, like that in Fig. 9.8. This equation assumes infinite DC-gain and error-free value of the each weight g_i . In a SC implementation of the integrator, the latter assumption amounts to saying that the amplifier has infinite DC-gain and that capacitors are perfectly matched – both impossible to achieve in practice.

Consider now the real case in which the amplifier DC-gain is high, but finite, and some mismatch is present. Analysis leads to the following modified finite-difference equation for the integrator,

$$v_{o,n} = \frac{A_V \cdot \sum_{i=1}^{n_b} g_i (1 - \epsilon_i)}{A_V + 1 + \sum_{i=1}^{n_b} g_i (1 - \epsilon_i)} v_{i,n-1} + \frac{(A_V + 1)}{A_V + 1 + \sum_{i=1}^{n_b} g_i (1 - \epsilon_i)} v_{o,n-1} \quad (9.8)$$

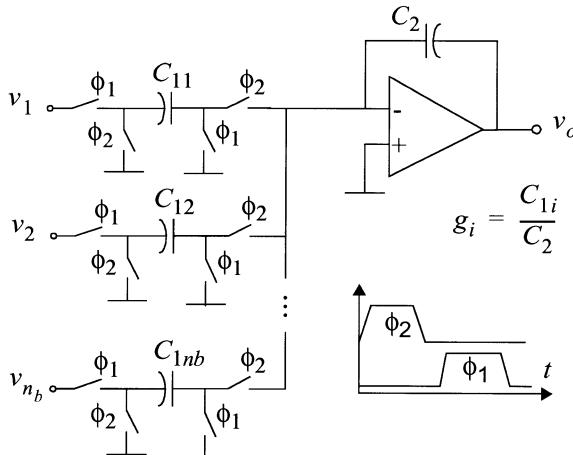


Figure 9.8: SC integrator with n_b input branches.

where A_V stands for the amplifier open-loop DC-gain, n_b is the number of input branches of the integrator, g_i is the nominal weight of the i -th branch, ε_i is its relative error, and $v_{i,n-1}$ and $v_{o,n-1}$ are the integrator input and output during the previous clock cycle. Apart from the evident error in the integrator weight, note that the coefficient of $v_{o,n-1}$ is smaller than unity. This effect, called leakage, attenuates the value of the integrator internal state from sample to sample.

When the real integrator expression is used instead of eq. (9.7) to re-compute (9.3), it is found that the quantization errors generated in stages other than the last one also appear at the modulator output. That is, there is a defective cancellation of the low-order quantization errors, caused by the leakage mechanism and the presence of mismatch between analog (integrator weights) and digital coefficients.

A detailed analysis of these effects is out of the scope of this chapter, but it can be found elsewhere [11]. Here we will make use of an approximate close expression for the extra quantization error power,

$$\Delta P_Q(A_V, \sigma_C) \cong \sigma_{sQ}^2 \left(\frac{25}{48} \frac{\pi^2}{A_V^2 M^3} + 24 \sigma_C^2 \frac{\pi^4}{5M^5} \right) \quad (9.9)$$

where $\sigma_{sQ}^2 = (2V_r)^2/12$ is the quantization error power of a single-bit quantizer and σ_C is the standard deviation of capacitor mismatch error. Although it has been approximated, eq. (9.9) can be applied to all the architectures covered using the set of integrator weights in eq. (9.4) and regular values of the oversampling ratio. When using other integrator weights, only small changes need to be entered in the numbers, but the dependence on M remains the same. This expression can also be used to evaluate other cascades, such as the 2-2 and 2-2-2, and whatever cascade starting with a second-order modulator. Qualitatively, this is explained by the fact that the most important contribution to the extra error power caused by integrator leakage is due to the first stage – term inversely proportional to M^3 in eq. (9.9)–, whereas that caused by coefficient mismatch is related to the link between the first and second stage - term inversely proportional to M^5 ^{†6}.

These contributions may indeed dominate the ideal one: (a) they are generated by single-bit quantizers, i.e. $\sigma_{sQ}^2 > \sigma_{mQ}^2$,^{†7}; (b), they are attenuated by only M^3 and M^5 . Hence, the feasibility of calibration-free cascade modulators depends on how demanding the requirements for the amplifier DC-gain and capacitor matching are. An evaluation of these concerns is given next.

First, we will identify a feasibility limit to the order of the single-bit architec-

6. If the first stage had been of first order, the leakage term would have exhibited a dependence with $1/M$ and the mismatching term with $1/M^3$, thus being more sensitive to these circuit imperfections. This is the reason why the first-order modulator was previously discarded as first stage of the cascade.

ture through behavioral simulation and MonteCarlo analysis [11] ^{†8}. Fig. 9.9(a) shows the simulated half-scale *SNDR* as a function of the amplifier DC-gain for $M = 16$. Fig. 9.9(b) shows the *SNDR* histograms obtained from MonteCarlo simulation assuming 0.1% sigma in capacitor ratios (0.05% is currently featured by MiM capacitors in CMOS processes [27]). Under these conditions, mainly because of the matching sensitivity, the 7th-order architecture is not worth implementing. Nevertheless, the 6th-order modulator provides 90-dB worst-case *SNDR* with DC-gain of 2500. Especially robust is the 5th-order cascade, which requires a DC-gain of 1000 to achieve 80-dB worst-case *SNDR* with $M = 16$.

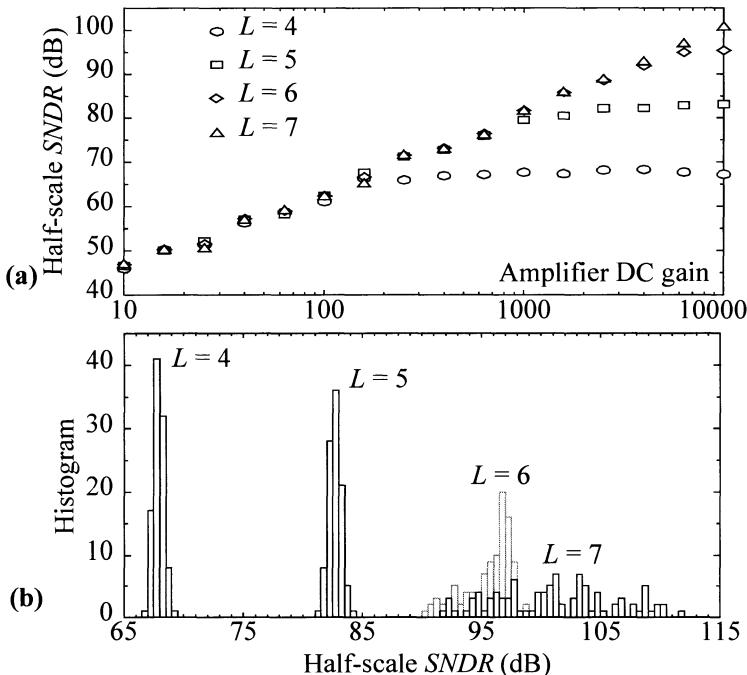


Figure 9.9: Effect of (a) finite DC-gain and (b) weight mismatch on the *SNDR* of single-bit $2-1^k$ modulators for $M = 16$.

-
7. Trying to alleviate this problem, some designers have adopted an all-multi-bit version of the cascade $\Sigma\Delta$ M, in which all the single-bit quantizers are replaced by multi-bit ones [6][7]. This strategy also relaxes the overloading issues, but at a price: the non-linearity of the multi-bit DAC in the first stage becomes critical and requires calibration/correction - precisely what we are trying to avoid.
 8. Simulation is preferred to calculation for its higher accuracy. Furthermore, the interested reader can verify the goodness of the analytical expressions above by comparing their predictions with those given by behavioral simulation directly implementing eq. (9.8).

In the same way as the modulator order, in practice the number of bits in the last-stage quantizer cannot be arbitrarily large. As shown in Fig.9.10, for a given oversampling ratio, the evolution of the effective resolution with B tends to saturate due to the presence of leakage and mismatching. In fact, when increasing the last quantizer resolution in 1bit does generate less than 1-bit increase in the overall resolution, the architecture starts losing efficiency and may become unsuitable for the specifications and technology considered. Nevertheless, depending on the signal bandwidth, the reduction in oversampling ratio that can be achieved by resorting to multi-bit quantization may define the border between feasible and infeasible implementations. For instance, consider again the fourth-order $2\text{-}1^2$ architecture. Now, according to Fig. 9.10, its single-bit version would require $M = 22$ to obtain 14-bit resolution, whereas a 3-bit version can achieve the same resolution with $M = 16$. For a signal bandwidth of 2.2MHz, these oversampling ratios mean 96.8MHz and 70.4MHz clock rate, respectively. Apart from an eased testing, certain power saving can be expected by using the multi-bit modulator. But this again depends on the characteristics of the fabrication process, as we explain in the next section.

9.3.2 Impact of deep-submicron features

The successful implementation of the cascade multi-bit architectures in incoming deep-submicron CMOS processes is strongly linked to two main technology features: (a) supply voltage and (b) capacitor performance.

The former is first related to the selection of the reference voltage used, which limits the available dynamic range. Second, the supply voltage also impacts the

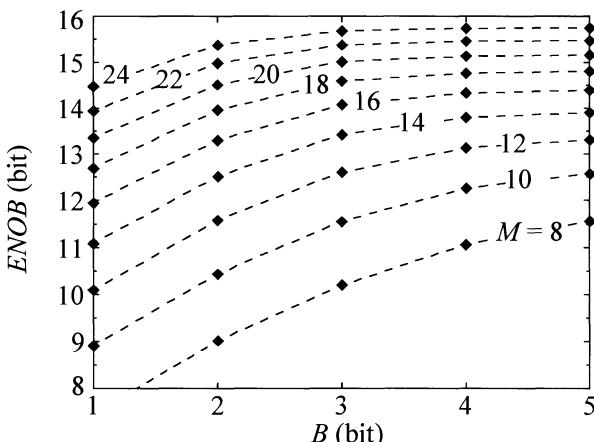


Figure 9.10: $ENOB$ vs. last-quantizer resolution for a $2\text{-}1^2mb$ $\Sigma\Delta M$ in the presence of circuit imperfections, $A_V = 2500$, $\sigma_C = 0.12\%$, $INL_{DAC} = 0.4\%FS$.

selection of the amplifier topology. In fact, two-stage amplifier topologies are required for some integrators in the cascade modulator for implementations in processes with 3.3-V supply and below. Definitely, this will be the choice in new CMOS technologies. The subsequent power penalty of the two-stage amplifier with respect to their single-stage counterparts will be the main bottleneck for optimizing the power/performance figure in future implementations of high-frequency $\Sigma\Delta$ Ms. We will return to the amplifier later on.

Apart from the amplifier optimization, the performance of the switches with supply voltages below 2.5V will have to be carefully surveyed. It is likely that the on-resistance linearity requirements will not be fulfilled by these switches, thus giving rise to limiting dynamic distortion. So, the employment of clock-boosting strategies will be mandatory, with the subsequent increase in circuit complexity and power dissipation. Another solution to this problem could be using higher voltage devices in those processes with double oxide thickness. However, when available, the use of such devices for signal processing is not recommended, probably for the lack of proper analog models. Low-VT devices are also optionally available in modern technologies, but at a higher price.

The second most relevant technology feature has to do with the quality of the capacitor structures. It has been shown that the capacitor matching requirements are in the range of 0.1% - 0.2% standard deviation. Low parasitics are also of extreme importance for an efficient implementation of a high-frequency modulator. Finally, we have the capacitor linearity requirements, which are less demanding provided that matched fully-differential circuitry is used. Fortunately, M-i-M (metal-insulator-metal) capacitor structures are now available in CMOS processes [27]. They exhibit an excellent matching and linearity, with very small bottom parasitics. If these capacitors are to be available in the new CMOS technology, they will not contribute any loss of performance.

In order to quantitatively evaluate the previous assumptions, we will develop an analytical procedure to roughly estimate the power consumption of different cascade single-bit and/or multi-bit sigma-delta modulators. In these expressions, both architecture and technological features will be contemplated, together with simplifying assumptions inspired in practical design solutions. The aim is not only to draw conclusions about architectural choices, but also to track their evolution under technology changes.

9.3.2.1 *Dominant error mechanisms in high-frequency $\Sigma\Delta$ modulators*

Let's start assuming that, in whatever practical design of a high-frequency $\Sigma\Delta$ modulator, the dominant sources of in-band error power are: quantization, circuit noise, and incomplete settling error. The latter is especially important for a telecom converter, in which sampling frequencies at the edge of the CMOS feasibility

will have to be used. In this section we will use simplified expressions for the in-band error power caused by these error mechanisms (following a procedure similar to those in [11][28]), to derive values for the main design parameters and, subsequently, an estimation of the power dissipation of a given cascade topology^{†9}.

Under the initial assumptions, the effective number of bits (*ENOB*) and the dynamic range (*DR*) of the modulator are related as follows,

$$DR = 3 \cdot 2^{2ENOB - 1} \equiv \frac{V_r^2 / 2}{P_Q + P_{Th} + P_{St}} \quad (9.10)$$

where V_r represents the full-scale input range of the modulator (which coincides with the reference voltage) and P_Q , P_{Th} , and P_{St} are the in-band powers of the quantization error, thermal noise, and incomplete settling error, respectively. At first glance, the right-hand side of eq. (9.10) states that the larger V_r , the better. However, the selection of the reference voltage also impacts P_Q and, although more indirectly, also P_{Th} and P_{St} . Moreover, V_r is obviously constrained by the supply voltage, because it imposes a given output swing requirement in integrators, which must be feasible in the technology considered. In conclusion, the selection of V_r is closely related to the type of amplifier forming the integrator and its capability to trade open-loop DC-gain, speed and output swing requirements^{†10}. In practice, an upper bound for a feasible selection of V_r is given by:

$$V_r = V_{\text{supply}} - n_{ob} V_{sat}, \text{ in volt-peak differential,} \quad (9.11)$$

where V_{sat} is the saturation voltage of the output devices and n_{ob} is the number of transistors in the output branch, which again depends on the specific amplifier topology. For the sake of simplicity, we will assume that cascode devices will be used in single-stage amplifiers. This means that our single-stage amplifier will be the popular folded-cascode OTA, with $n_{ob} = 4$. However, as stated before, this popular choice is not always enough to achieve a good DC-gain, output swing trade-off. This is specially true for low supply voltage implementations, where an excessive value of V_{sat} will result in a ridiculously small value for the feasible reference voltage, thus producing little efficient modulators.

9. The reader is warned of the empirical nature of some of the expressions in this section. They are based either on our own design experience or are commonly found in the open literature. So, the employment of these expressions will be just briefly justified.
10. This triple trade-off among gain, speed and output swing is always present in an amplifier. A high open-loop gain requires that the transistors are deep in saturation, which forces to use large V_{DS} voltages, thus reducing the available voltage swing. This effect is worsened when a large device current is required for dynamic considerations, because the saturation voltage is increased for constant dimensions [29][30].

Among the alternatives, we count on two-stage amplifiers [29][30], whose output branch can contain only two transistors ($n_{ob} = 2$), still producing a large open-loop DC gain. This allows us to increase the value of the reference voltage up to useful levels. For example, with two-stage amplifiers and assuming 0.25-V saturation voltage, a differential reference voltage of 2V would fit in a 2.5-V supply technology, almost the same as in a 3.3-V supply process using single-stage folded-cascode OTAs.

Next, following our simplified approach, we will assume for the time being that the incomplete settling error can be controlled by design so that $P_{St} \ll P_Q, P_{Th}$; that is, eq. (9.10) simplifies to:

$$DR \approx \frac{V_r^2 / 2}{P_Q + P_{Th}} \quad (9.12)$$

With respect to P_Q , it is formed by 3 main error mechanisms: (a) last-stage quantization error; (b) last-stage DAC non-linearity (for multi-bit quantization only); and (c) non-cancelled portion of the low-order quantization errors caused by integrator leakage and mismatching among analog and digital coefficients. A close expression, including all the non-idealities above, can be derived from those in Section 9.3.1, resulting in

$$\begin{aligned} P_Q = & \sigma_{mQ}^2 d^2 \frac{\pi^{2L}}{(2L+1)M^{2L+1}} + \sigma_{INL}^2 d^2 \frac{\pi^{2(L-1)}}{(2L-1)M^{2L-1}} \\ & + \sigma_{sQ}^2 \left(\frac{25}{48} \frac{\pi^2}{A_v^2 M^3} + 24 \sigma_C^2 \frac{\pi^4}{5M^5} \right) \end{aligned} \quad (9.13)$$

where the three terms stand for contributions (a), (b) and (c), respectively. Notation in the previous equation has already been defined in Section 9.3.1. Here we repeat the values of the error powers:

$$\sigma_{mQ}^2 = \left(\frac{2V_r}{2^B - 1} \right)^2 / 12 \quad \sigma_{INL}^2 = \frac{(2V_r \times INL)^2}{2} \quad \sigma_{sQ}^2 = \frac{(2V_r)^2}{12} \quad (9.14)$$

The second contribution to the simplified in-band error power in eq. (9.12) is circuit noise, P_{Th} . In our case, the circuit noise power will be usually dominated by thermal noise. Other noise sources, such as $1/f$ noise, are intended for playing a secondary role, the reason being twofold: (a) DC and the low-frequency region of the spectrum are normally out of the signal band in telecom applications; (b) the usage of small capacitors in order to relax the dynamic requirements increases kT/C noise over the flicker noise components^{†11}. A conservative expression for the in-band power of thermal noise can be derived [11],

$$P_{Th} \cong \frac{16kT}{3MC_s} \quad (9.15)$$

where C_s is the value of the sampling capacitor.

9.3.2.2 Static power consumption

By going over eqs. (9.10) to (9.15), it is found that the dynamic range of a cascade $\Sigma\Delta$ modulator can be roughly expressed as a function of the following design parameters: V_{supply} , L , M , C_s , A_v , and σ_C , to which we have to add B and INL if the last-stage quantizer is multi-bit. So, for given values of A_v , σ_C , and INL , previous equations allow us to calculate the minimum value of the capacitor C_s required to obtain a given DR , as a function of M , L , and B . Once C_s is known, the equivalent load for the amplifier in the integrator can be estimated as,

$$C_{eq} \cong C_s + C_p + C_l \left(1 + \frac{C_s + C_p}{C_o} \right) \quad (9.16)$$

where C_o , the integrator feedback capacitance, is related to C_s through the integrator weight, $C_o = C_s/g_i$; and C_p , C_l stand for the integrator summing node and output parasitics, respectively. Estimating the latter two capacitances is a difficult task because of their extreme dependence on the actual amplifier design^{†12}.

Usually, the main contribution to C_p is the amplifier input parasitics. In a fully-differential topology, this is formed by the input transistor gate-to-source capacitance C_{gs} (both channel and overlap contributions) and its overlap gate-to-drain capacitance C_{gd}^{ov} , amplified by Miller effect. Thus,

11. Of course, the veracity of this assumption depends on how large the $1/f$ contribution is, and where its corner frequency is located. For very small MOS devices flicker noise can be huge [31][32], thus dominating the in-band error power in deep-submicron implementations. Apart from this consideration, device mismatching and, consequently, worsened sensitivity to common-mode interferences precludes designers from making use of minimum length MOS devices. Moreover, it is generally true that the matching properties of minimum-length devices is getting better and better as the technology shrinks, whereas their $1/f$ noise, especially for n-channel MOSFETs, increases. If this trend goes on in the coming technologies, flicker noise may become the most important consideration for selecting the transistor length, even in high-frequency applications.

12. Apart from a small part due to interconnects, the input capacitance C_p is mainly contributed by the switches and the amplifier input stage. We will assume that the latter is dominant. The same can be said with respect to the output parasitics C_l , to which the contribution of the bottom-plate parasitics of the integration capacitor is normally smaller than the amplifier output parasitic capacitance. In any case, the reader will understand that the finest possible estimations should be pursued for C_p and C_l , given their impact in the integrator dynamics and, hence, in power dissipation.

$$C_p \equiv C_{gs}^{\text{ch}} + C_{gs}^{\text{ov}} + C_{gd}^{\text{ov}}(1 + A_{v_1}) = \frac{2}{3}C_{ox}'W_{in}L_{in} + C_{ox}'W_{in}\Delta L_{in}(A_{v_1} + 2) \quad (9.17)$$

where C_{ox}' is the gate oxide capacitance density, and ΔL_{in} stands for the lateral diffusion of drain/source regions below the gate, both technology-dependent parameters^{†13}. Apart from the input transistor dimensions W_{in}, L_{in} , the other unknown variable in eq. (9.17) is its input-to-output gain A_{v_1} . This is equal to the complete gain of the amplifier for single-stage amplifiers, or to the first-stage gain if multi-stage topologies are used. It can even be around unity if cascode devices are used, such as in folded- or telescope-cascode OTAs [29][30]. Now, let's make use of the expression for the input transistor drain current, which in strong inversion saturation can be approximated to:^{†14}

$$I_{D,in} \approx \frac{\mu C_{ox}' W_{in}}{2 L_{in}} (V_{GS} - V_T)^2 \quad (9.18)$$

Using (9.18) to replace the product $C_{ox}' W_{in}$ in eq. (9.17) results in

$$C_p = \frac{2L_{in}I_{D,in}}{\mu V_{OVD}^2} \left[\frac{2}{3}L_{in} + \Delta L_{in}(A_{v_1} + 2) \right] \quad (9.19)$$

where $V_{OVD} \equiv V_{GS} - V_T$ is the input transistor overdrive voltage.

With respect to C_l in eq. (9.16), we assume that the main contribution is due to the amplifier output parasitics. So, it depends on the actual sizes of the devices connected to the output node. The value of C_l is influenced by many features, for example: whether the amplifier is single- or multi-stage, the output stage contains or not cascode devices, etc. Even the supply voltage, via output swing and DC-gain requirements, makes an impact on the transistor sizes and hence on C_l . All things considered, a reliable estimation of this capacitance prior to sizing the actual amplifier is not possible. Since whatever estimate could in practice be 100% erroneous, let us adopt the simplest possible assuming that C_l is constant.

13. In estimating C_p , the gate-to-substrate capacitance $C_{gb} = C_{gb}^{\text{ch}} + C_{gb}^{\text{ov}}$ has been neglected.

On the one hand, the device normally operates in strong inversion, so that $C_{gb}^{\text{ch}} \approx 0$; on the other, the overlap contribution C_{gb}^{ov} will be negligible as compared to $C_{gd}^{\text{ov}}, C_{gs}^{\text{ov}}$ because usually $W_{in} \gg L_{in}$ in high-frequency amplifiers [31].

14. This expression can be rather inaccurate, mainly for short-channel devices [31][32]. On the one hand, a large mobility degradation is expected, making the current decrease with respect to eq. (9.18); on the other, a normally large output conductance will make the current increase. For the sake of simplicity, we can suppose that both effects cancel each other, to first-order approximation. However, the impact of the temperature on the mobility should be included. To this purpose, we can use the approximate relation $\mu(T) \approx \mu(T_0)(T/T_0)^{-2}$, where T_0 is the room temperature [31].

This coarse approach is, nevertheless, reasonable to some extent. Let's take the case of an amplifier which is going to be migrated from one technology to another, the latter being smaller and using lower supply voltage. Scaled dimensions in the new technology would result in smaller parasitics, but the shrunk supply voltage requires higher aspect ratios to accommodate the same output swing, which, in its turn, would increase parasitics. Thus, both trends result in an essentially unchanged parasitic output load^{†15}.

To be consequent with the initial assumption $P_{St} \ll P_Q, P_{Th}$, the incomplete settling error power must be small enough to obtain the required resolution. However, an accurate estimation of the settling error would involve the following calculations. For example, just for a single-pole amplifier model, complicate expressions are derived [11] if a non-linear (slew-rate limited) settling is considered. Further complexity arises from considering both sampling and integration incomplete charge-transference [33] and the contribution of the non-zero switch on-resistance. Hence, we will simplify our treatment assuming that the slew-rate of the amplifier is large enough to neglect its impact on the integrator transient response, so that the settling is linear with time constant equal to C_{eq}/g_m . This being the case, it takes a number $\ln(2^{ENOB})$ of time constants to settle within $ENOB$ resolution; that is, the following relation should be fulfilled:

$$\ln[2^{(ENOB+1)}] \frac{C_{eq}}{g_m} \leq \frac{T_S}{2} \quad (9.20)$$

where T_S is the sampling period. Note that we have added an extra bit in order to make room for other error mechanisms, and account for the inaccuracy of this simplified model. The above expression can be used to estimate the minimum value of the transconductance parameter as,

$$g_m = 2f_S \ln[2^{(ENOB+1)}] C_{eq} \quad (9.21)$$

where $f_S \equiv 1/T_S$ is the sampling frequency. This is the transconductance required for a single-stage amplifier, for which C_{eq} in eq. (9.16) is the equivalent output load. For multi-stage amplifiers, the previous relation must be carefully tackled because both parameters, total transconductance and equivalent output load, lose

15. A more careful analysis shows that C_l tends to increase with technology scaling. Straight forward calculation shows that the new technology transistor width can be estimated as $W_2 \approx W_1(V_{DD,1}/V_{DD,2})^2(L_2/L_1)$, where subscripts 1 and 2 refer to the old and new process, respectively. Due to the presence of the squared ratio of the supply voltages, the ratio W_2/W_1 will normally exceed unity. In addition, the overlap capacitance per width unit (CGDO, CGSO) remains almost constant from process to process, so that C_l would increase.

control of the amplifier dynamics. However, provided that the main pole of the amplifier is set by the input stage and an eventual inter-stage compensation capacitor, eq. (9.21) can still be used to determine the input stage transconductance, which can be related to the input transistor current using the well-known approximate expression^{†16},

$$g_m = \frac{2I_{D,in}}{V_{OVD}} \quad (9.22)$$

Equations (9.16), (9.19), (9.21), and (9.22) can be handled in an iterative manner to determine the current required through the input transistors of the amplifier.

At this point one more assumption is necessary concerning the topology of the amplifier. Whereas for 5-V supply technologies the most employed amplifier topology has been the well-known single-stage folded-cascode OTA, two-stage amplifiers are gaining ground as technologies scale down and supply voltages shrink. Moreover, in practice two gain stages are not enough to achieve the overall gain requirement, so that the first one often includes cascode devices^{†17}.

Let's consider this topology as an archetype in modern deep submicron technologies and try to estimate its power dissipation. The current through the first stage has been already estimated as $2I_{D,in}$. Determining the current through the second stage is again hard. Several specifications such as slew-rate, phase margin, DC-gain, gain-bandwidth product, etc., are involved, rendering the analytical problem intractable. So, we have no alternative but to assume a fixed ratio between the current through the second and first stages, so that the total current through the amplifier could be,

$$I_B \cong 2I_{D,in} + 2\eta_{io}I_{D,in} + I_{D,in} = [2(1 + \eta_{io}) + 1]I_{D,in} \quad (9.23)$$

where η_{io} stands for the ratio $I_{D,out}/I_{D,in}$, and an extra $I_{D,in}$ has been added to account for the current used in the amplifier biasing stage^{†18}.

With eq. (9.23) the power dissipation of the first amplifier can be estimated. That of the remaining amplifiers in the cascade stages can be decreased with respect to the former, following the scaling rule commonly applied to the amplifier requirements in $\Sigma\Delta$ Ms [3]-[10]. This power reduction may come from either a

16. This expression can be much more inaccurate than eq. (9.18), from which it is derived, since a good approximation to the value of the current does not imply a good approximation to its derivative.

17. Usually in a telescope-cascode configuration [29].

18. A sound choice is $\eta_{io} = 3$. The factor 2 in eq. (9.23) comes from the fact that a fully-differential implementation is assumed. We have intentionally omitted the current dissipation of the common-mode feedback net because, normally, a switched-capacitor net will be used, whose power consumption is negligible.

relaxed set of specifications, or the subsequent amplifier topology simplification^{†19}. For the architectures covered in this chapter, we can write

$$I_{B, total} = I_B \left(1 + \sum_{i=2}^L \chi_i \right) \quad (9.24)$$

where χ_i is the ratio between the current absorption of the i -th amplifier and the first one. From this, the static power dissipated in amplifiers amounts to,

$$P_{op, sta} = I_B V_{\text{supply}} \left(1 + \sum_{i=2}^L \chi_i \right) \quad (9.25)$$

Some extra static power would be dissipated in an eventual preamplifying stage preceding the latched comparator of each stage, in a single-bit case. In practice the static current through each comparator is very small as compared with the one in amplifiers, causing little impact on the overall static power dissipation. Nevertheless, in a multi-bit case, the total number of comparators must be increased to implement the multi-bit last-stage quantization, thus increasing its relative power dissipation. So, certain power per preamplified latch, $P_{comp, sta}$, must be added to eq. (9.25).

Also in the multi-bit case, the last-stage multi-bit DAC contribution must be considered. Looking for the simplest possible DAC circuitry, this block is implemented through a resistor ladder^{†20}, which dissipates static power. The amount of current through the resistors must be fixed, according to dynamic requirements, in order to provide a good settling for the multi-bit reference voltages. Note that the current requirement scales with the sampling frequency and the capacitive load involved. Let's take the first scaling rule to write:

$$P_{DAC, sta} \cong V_{\text{supply}} I_{DAC, ref} \times \frac{f_S}{f_{S, ref}} \quad (9.26)$$

In this empirical expression, $I_{DAC, ref}$ is the current through the resistor ladder DAC required for operating at a certain frequency of reference, $f_{S, ref}$.

All things considered, the total static power dissipation can be estimated as,

$$P_{sta} = P_{op, sta} + P_{DAC, sta} + [(L-1) + (2^B - 1)] P_{comp, sta} \quad (9.27)$$

19.In fact, a single-stage topology could be used for amplifiers other than the first one, even in those cases in which a two-stage front-end amplifier is mandatory.

20.The same ladder can be used for generating the reference voltages for the ADC, so that some static power can be saved.

where the quantity in square brackets is the number of comparators in a L th-order cascade with B bits in the last-stage quantizer.

9.3.2.3 Dynamic power consumption

Although usually smaller than its static counterpart, the dynamic power dissipation should be taken into account, specially for high-resolution high-speed applications. The reasoning behind this statement becomes clear if we consider that the dynamic power dissipated to switch a capacitor of value C_u between the reference voltages at a frequency f_S can be estimated as,

$$P_{\text{dyn}} = C_u f_S V_r^2 \quad (9.28)$$

Note the influence of the three design parameters, and the fact that they all need to be large in a high-resolution, high-speed scenario^{†21}, so that the relative weight of the dynamic power tends to increase for these type of specifications. Its absolute value depends on the number and size of capacitors used in each integrator, which is ultimately related to the value of the integrator weights. For the integrator weights in eq. (9.4), the following formula can be used for dynamic power estimation,

$$P_{C, \text{dyn}} = 2 \times [5C_{u_1} + 4(L-1)C_{u_2}]f_S V_r^2 \equiv [10 + 8(L-1)\eta_{C_u}]C_{u_1} f_S V_r^2 \quad (9.29)$$

where the factor 2 comes from the differential implementation; C_{u_1} is the unitary capacitor used in the first integrator, whereas C_{u_2} is the one used in the second and following integrators. The latter is usually set to a smaller value than the former, the reason being that its impact on thermal noise is attenuated by increasing powers of the oversampling ratio^{†22}; so that $C_{u_2} = \eta_{C_u} C_{u_1}$, with $\eta_{C_u} < 1$.

In addition to this dynamic consumption, the dynamic power dissipated in the digital part of the modulator (quantizers, flip-flops and gates, cancellation logic, etc.) should be considered. Apart from being small, only minor differences are expected among the architectures considered, so that this dynamic power contribution can be neglected for comparison purposes^{†23}.

9.3.3 Comparison among cascade architectures

The expressions above can be used for comparing the performance of different cascade architectures. Here, this comparison is made from a twofold perspective:

-
- 21. Justifications for V_r and f_S are straight forward. In addition, the value of C_u must be large enough to provide a margin with respect to thermal noise.
 - 22. While thermal noise is the usual criterion for selecting the first integrator unitary capacitor, those of the second and following integrators are normally set trading capacitor matching and integrator dynamics considerations.

(a) for a given technology, the performance of several modulators are compared for varying converter specifications; and (b) for given converter specifications, the architectures are evaluated in the technology road map. In both cases, the following figure-of-merit (*FOM*)^{†24} has been used,

$$FOM = \frac{Power}{2^{ENOB} \times DOR} \times 10^{12} \quad (9.30)$$

where *DOR* stands for the digital output rate, i.e., the Nyquist rate. This *FOM* has been estimated using the power equations of the previous section.

As stated in Section 9.2, there are three design parameters univocally describing the cascade architectures covered in this chapter. They are: (a) modulator order, *L*, (b) oversampling ratio, *M*, and (c) number of bits of the last-stage quantization, *B*. Each triad {*L*, *M*, *B*} codifies an architecture.

In a first comparison task, these triads have been evaluated along the curve in the resolution - speed plane shown in Fig. 9.11 (dashed line). Although, this particular resolution - speed relationship is arbitrary, it fits the usual requirements for wireline telecom ADCs: ISDN, ADSL, VDSL, etc., which have been placed in the figure for illustration. The evaluation procedure consisted of the following steps:

- Selection of the digital output rate. The corresponding effective resolution, in terms of *ENOB*, is given by the resolution - speed relationship imposed.
- Selection of the technology. In our simplified approach, it is characterized by only four parameters: minimal transistor length, supply voltage, capacitor mismatch (σ_C), and resistor mismatch. The latter data, together with a prediction of the amplifier DC-gain, A_V ,^{†25} are needed for computing the non-ideal version of the quantization error power in eq. (9.13). Note that the resistor mismatch is related to the integral non-linearity of the resistive-ladder DAC and, hence, the value of σ_{INL} in eq. (9.14).
- Setting of other design parameters, such us maximum temperature, required for thermal noise computation, and the value of the overdrive voltage,

23. Note that, although it cannot be neglected at all, the dynamic power consumption of a digital filter following the modulator plays a similar role, not showing large differences for the architectures considered. The order of the digital filter must be at least that of the modulator plus one. So the higher the latter, the more complex the filter must be. However, an increase of the modulator order entails a decrease of the oversampling ratio and hence the filter can be operated at a lower frequency.

24. This *FOM*, a measure of the energy needed per conversion in pJ, has been extensively used for evaluating A/D converters [34].

25. The prediction about the open-loop DC-gain must be in accordance with the type of OTA selected. For example, if single-stage OTAs will be used, then $A_V \sim 500 - 1000$; if these are folded-cascode, then $A_V \sim 1000 - 2000$; finally, if two-stage OTAs are adopted, A_V can be much larger, say over 5000. These numbers decrease as technologies shrink.

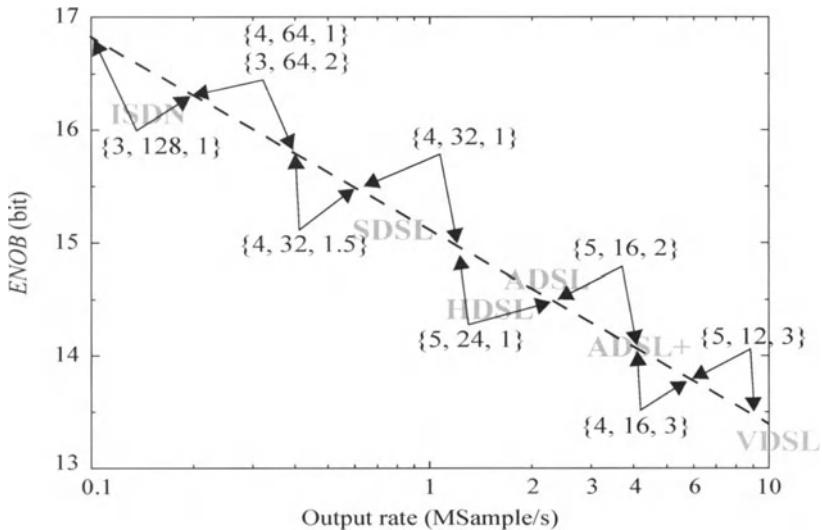


Figure 9.11: Recommended cascade architecture for each section of the resolution - speed plane, in a 2.5-V 0.25 μ m CMOS technology. Two-stage OTAs with $n_b = 2$ have been assumed. Other parameters were: $\sigma_C = 0,1\%$, $INL = 0,25\%$ FS, $A_V = 3000$, temperature = 110°C, and $V_{OVD} = 0,25$ V.

- Estimation of the power consumption of all possible triads $\{L, M, B\}$, which are ranked by the value of their FOM .

Fig. 9.11 shows the results for the conditions listed in the figure caption. For each section of the resolution-speed curve, the architecture with the minimum FOM has been noted down. Observe that the oversampling ratio decreases as the output rate increases and, simultaneously, the multi-bit quantization shows up to compensate for the oversampling reduction. In order to make evident the impact on power dissipation of a wrong architectural choice, some of the triads together with their relative FOM are listed in Table 9.1, assuming the previous conditions, and looking for 14bit at 4.4MS/s. Note that the fifth-order single-bit cascade consumes 11% more than the fourth-order 3-bit modulator. Surprisingly, some popular cascades, such as {3, 24, 3} and {4, 24, 1}, are significantly less efficient for the specifications and technology considered.

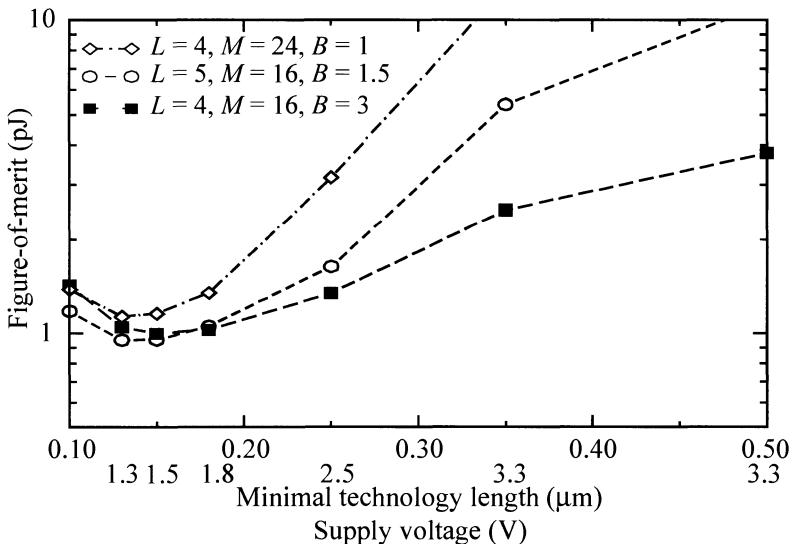
This is leading us to the other perspective of the architecture exploration. The goal now is to know how the performance of the cascade multi-bit architectures is going to evolve under technological changes. Since some technology features enter the approximate formulation in the previous section, it is possible to use these expressions to track the approximate power dissipation of a given topology

Table 9.1: Architecture rank for 14bit@4.4MS/s in a $0.25\mu\text{m}$ CMOS process.

Rank	$\{L, M, B\}$	Relative <i>FOM</i>
1	$\{4, 16, 3\}$	1.000
2	$\{5, 16, 1\}$	1.108
...
6	$\{3, 24, 3\}$	1.538
7	$\{4, 24, 1\}$	2.590

and set of modulator specifications. For example, Fig. 9.12 shows the estimated evolution of the *FOM* of three cascade topologies, namely $\{4, 24, 1\}$, $\{5, 16, 1.5\}$ and $\{4, 16, 3\}$, aimed at obtaining 14bit at 4.4MS/s. Two facts are noticeable:

(a) Despite the reduction of the supply voltage, overall, the power dissipation does not decrease below $0.18\mu\text{m}$. It even increases for most architectures²⁶. This is basically due to the reduction in supply voltages, which imposes a reduction in

Figure 9.12: Estimated evolution of the *FOM* with technology for three cascade architectures obtaining 14bit@4.4MS/s.

26. The location of the inflection point is optimistic because at $0.18\mu\text{m}$ and below, the switch non-linearity will probably dominate (depending on the converter requirements). On the one hand, this fact will reduce the dynamic range; on the other, some extra power will be needed for attenuating its effect through the use of clock boosting strategies, for example.

the reference voltage and, hence, the signal power. In order to keep the effective resolution, this loss of signal power must be compensated by an equivalent reduction of the noise power, which can be achieved by increasing the value of the sampling capacitors. Since the incomplete settling error power must be also kept constant, the larger the capacitances, the more demanding the dynamic requirements for the amplifiers. Whereas up to $0.18\mu\text{m}$ the increase in current absorption caused by this mechanism is compensated, in terms of power, by the supply voltage scaling, the estimated trend is exactly the opposite below that technology. Once again, the location of the inflection point depends on the converter specifications. For instance, if for the same speed the resolution is to be increased, the inflection point moves to the right in Fig. 9.12.

(b) Another aspect illustrated in Fig. 9.12 is the dynamic nature of the architecture selection in Fig. 9.11. Note the evolution of the $\{4, 16, 3\}$ $\Sigma\Delta M$. It outperforms for $0.25\mu\text{m}$ and above, but it does not below $0.18\mu\text{m}$. The reason behind this is that the multi-bit modulator has a fixed amount of power contributed by the last-stage quantizer that is not present in the single-bit implementation ($\{4, 24, 1\}$ $\Sigma\Delta M$). In addition, the latter takes advantage of the faster technologies to compensate for the increased oversampling ratio with respect to the multi-bit modulator. A special mention requires the case of the fifth-order $\{5, 16, 1.5\}$ $\Sigma\Delta M$. If the quality of the M - i - M capacitors is preserved, this architecture will be worth exploring in the coming technologies.

9.4 DESIGN EXAMPLE

In order to illustrate the previous considerations, we face the design of a $\Sigma\Delta M$ with 14-bit effective resolution at 4.4MS/s in a 1.8-V $0.18\mu\text{m}$ CMOS technology. These specifications, required for ADSL+ modems [1], pose a significant design challenge: On the one hand, both resolution and speed are at the state-of-the-art top edge. On the other, high-resolution must be attained employing a single 1.8-V supply.

After applying the topology selection method described in Section 9.3.3, the $2^{12}mb$ architecture with $M = 16$ and $B = 3$ (that is, the $\{4, 16, 3\}$ modulator) turned out to be the lowest-FOM solution, with 63-mW power estimate. Apart from L , M , and B , preliminary values for the sampling capacitor and the reference voltage are also set in this initial step: $C_s = 1.37\text{pF}$, $V_r = 1.3V_{pd}$.

9.4.1 Switched-capacitor implementation

Fig.9.13 shows the fully-differential SC schematic of the modulator. The first stage of the modulator includes two SC integrators, the first one with a single input branch and the second with two input branches. Switches controlled by the

comparator outputs are employed to feed back the quantized signal. The second stage incorporates an integrator with only two input branches, although three different weights are implemented: g_3 , g'_3 , and g''_3 . This can be done because the selection of weights allows distribution of the weight $g_3 = 1$ between the two

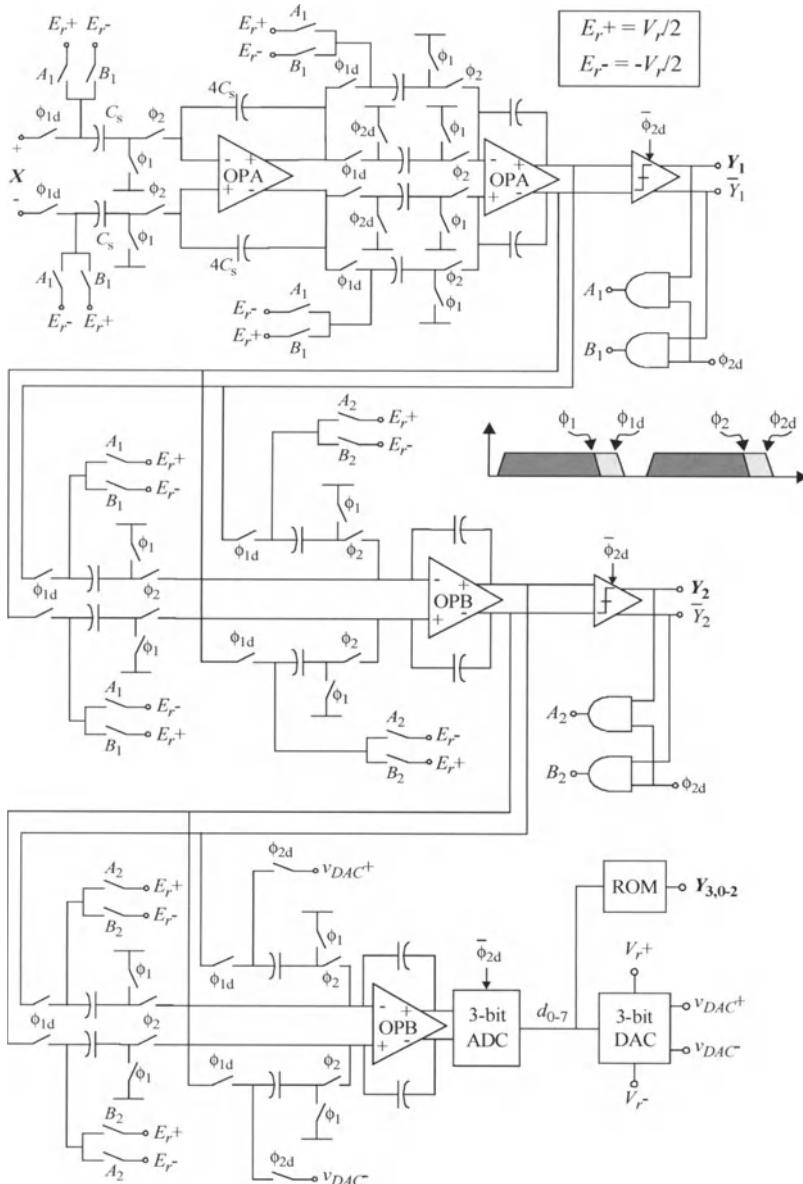


Figure 9.13: SC implementation of the $\{4, 16, 3\}$ $\Sigma\Delta M$.

integrator branches: $g'_3 = 0,5$ and $g''_3 = 0,5$. The same applies for g_4 in the fourth integrator of the cascade, also requiring only two input branches. Note that the integrator weights of the fourth-stage have been scaled with respect to those in eq. (9.4) in order to keep a loop gain equal to unity in the multi-bit stage, so that $g_4 = 2$, $g'_4 = g''_4 = 1$. Nevertheless, this change causes no impact on other characteristics. The last integrator drives the 3-bit ADC and the third-stage loop is closed through a 3-bit DAC.

The modulator operation is controlled by two non-overlapped clock-phases. In order to attenuate the signal-dependent clock-feedthrough, delayed versions of the two phases, ϕ_{1d} and ϕ_{2d} , are also provided. As illustrated in Fig. 9.13, this delay is incorporated only to the falling edges of the clock-phases, i.e., to the turn-off of the switches, while the rising edges are synchronized in order to increase the effective time-slot for the modulator operations [35]. The comparators and the ADC are activated at the end of phase ϕ_2 , using ϕ_{2d} as a strobe signal, to avoid any possible interference due to the transient response of the integrators outputs in the beginning of the sampling phase [3].

9.4.2 Building block specifications

The modulator specifications are mapped onto requirements for its building blocks using SDOPT (an equation-based optimizer) and ASIDES (a time-domain behavioral simulator) [11]. The main issues covered in these tools are integrator/amplifier and quantizer errors, passive component errors, analog switch errors, mismatch, thermal noise, etc. Particularly, the tools have been recently updated with new, accurate models for capacitor mismatch, thermal noise, and settling behavior of SC integrators [33].

Table 9.2 summarizes the sizing results for the modulator achieving 14bit@4.4MS/s. Note that the main in-band error source is due to quantization noise, which amounts to -89dB , while other error mechanisms are well below the limit imposed by quantization error except for circuit noise. This is because, given the high-resolution, high-speed nature of this application, the choice of the sampling capacitor, involved in both kT/C noise and dynamic issues, is critical. Taking into account this trade-off, the tool selected a value of 1.25pF , which is slightly smaller than the preliminary value (1.37pF); the reason being that the equations handled at this level are more accurate than those used for power estimation^{†27}. The output swing requirement (OS) deserves special attention in a 1.8-V supply implementation. Thanks to the previous selection of integrator weights, the OS can be relaxed to only $\pm 1,3\text{V}$, which is feasible in a differential approach.

27.In our design methodology the accuracy of the models increases as we approach the electrical/physical level.

The amplifier dynamics has been mapped onto its gain-bandwidth (GBW) product, phase margin (PM), and slew-rate (SR). The corresponding specifications are accompanied by the equivalent output load. Note that this is different for GBW and SR , the reason being that the amplifier equivalent load (in the integrator) actually changes from linear to slewing conditions.

The previous integrator/amplifier specifications apply to the first integrator. Some specs, such as DC-gain and dynamic performance, can be relaxed for the remaining integrators, because their in-band error contributions are attenuated by increasing powers of the oversampling ratio. For the same reason, also the sampling capacitor value can be reduced with respect to the value used in the front-end integrator. The fact that the modulator exhibits less sensitivity to non-idealities associated to the integrators located at the back-end of the cascade provides a strategy for reducing power dissipation, consisting in sizing a dedicated amplifier

Table 9.2: Sizing results for the {4, 16, 3} $\Sigma\Delta M$.

SPECS: 14bit@4.4MS/s		2-1 ² <i>mb</i>	Unit
MODULATOR	Oversampling ratio	16	
	Multi-bit quantization	3	bit
	Sampling frequency	70.4	MHz
	Reference voltages	± 1.3	V
INTEGRATORS	Unitary capacitor	1.25	pF
	Mismatch standard deviation	0.1	%
	Capacitor non-linearity	15	ppm/V
	Bottom parasitic capacitor	1	%
	Switch on-resistance	200	Ω
AMPLIFIERS	Open-loop DC-gain	2700	
	DC-gain non-linearity	10%	V^{-2}
	$GBW (C_L = 5.2\text{pF})$	220	MHz
	$PM (C_L = 5.2\text{pF})$	60	$^\circ$
	$SR (C_L = 3.84\text{pF})$	481	$V/\mu\text{s}$
	Input eqv. thermal noise PSD	3	$\text{nV}/(\text{Hz})^{1/2}$
	Input parasitic capacitance	0.5	pF
	Differential OS	± 1.3	V
COMPARATORS	Hysteresis	20	mV
A/D/A CON- VERTER	Resolution	3	bit
	Integral non-linearity	0.25	%FS
Dynamic range		86.1dB (14.0bit)	
Quantization noise		-89.0dB	
Circuit noise (KT/C + amplifier noise)		-91.1dB	
Incomplete settling noise		-103.1dB	
Harmonic distortion		-110.9dB	

Table 9.3: Fine-tuning of the amplifier specifications.

SDOPT	ASIDES			
	1st intg.	2nd intg.	3rd intg.	4th intg.
Unitary capacitor	1.25pF	1.25pF	0.3pF	0.3pF
GBW (5.2pF)	220MHz	190MHz	145MHz	
SR (3.84pF)	481V/ μ s	442V/ μ s	382V/ μ s	
DC-gain	2700	2600	780	
Differential OS	$\pm 1.3V$	$\pm 1.3V$	$\pm 1.3V$	

for each integrator. As an intermediate solution, more practical from the electrical design point of view, only two different OTAs will be considered in this example. Their specifications are obtained through fine-tuning at the modulator level with the help of ASIDES. Results are in Table 9.3. Note that, overall, the OTA specifications have been relaxed. This is especially true for the DC-gain of the second OTA. and for its dynamics. As shown in Fig. 9.14, this reduction seems to jeopardize the third and fourth integrator transient responses. However, the low sensitivity of the modulator to these errors allows us to obtain full performance. Fig. 9.15 shows the final output spectrum and $SNDR$ curve obtained with ASIDES.

9.4.3 Building block implementation

Fig. 9.16(a) shows a 2-stage 2-path compensated OTA, with a telescopic-cascade first-stage and both Miller and Ahuja compensation. The second stage input

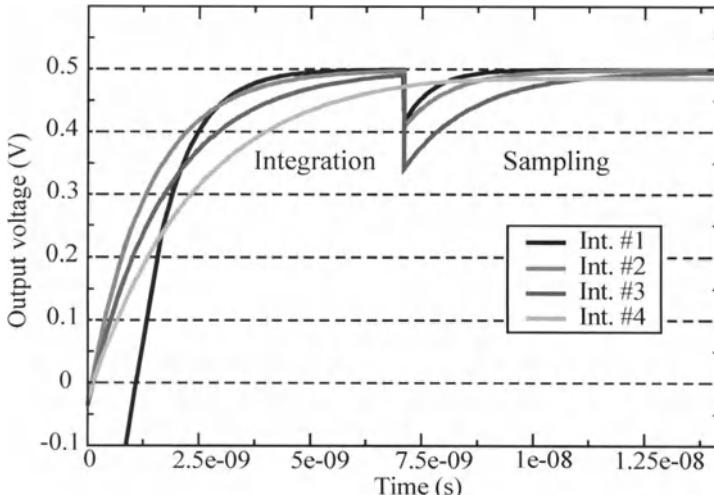


Figure 9.14: Integrator transient response to an input step that causes an increase of 0.5V at each integrator output (obtained from ASIDES).

is actually a differential pair, which allows excellent control of its current with almost rail-to-rail *OS*. Note that in the first stage pMOS input transistors have been preferred to nMOS because: first, the latter usually contribute more $1/f$ noise, and second, we have assumed that substrate is p-type, so that it is possible to remove the body-effect of the pMOS devices, thus eliminating one of the paths for substrate noise coupling. However, the second stage uses nMOS input devices to exploit their larger transconductance. Two version of this amplifier have been designed using FRIDGE [11]: one for the first and second integrators (OPA), and another for the third, and fourth integrators (OPB). Their simulated performances are summarized in Table 9.4. Additionally, in low-voltage implementations the impact of the OTA DC-gain non-linearity must be carefully checked. Fig. 9.17

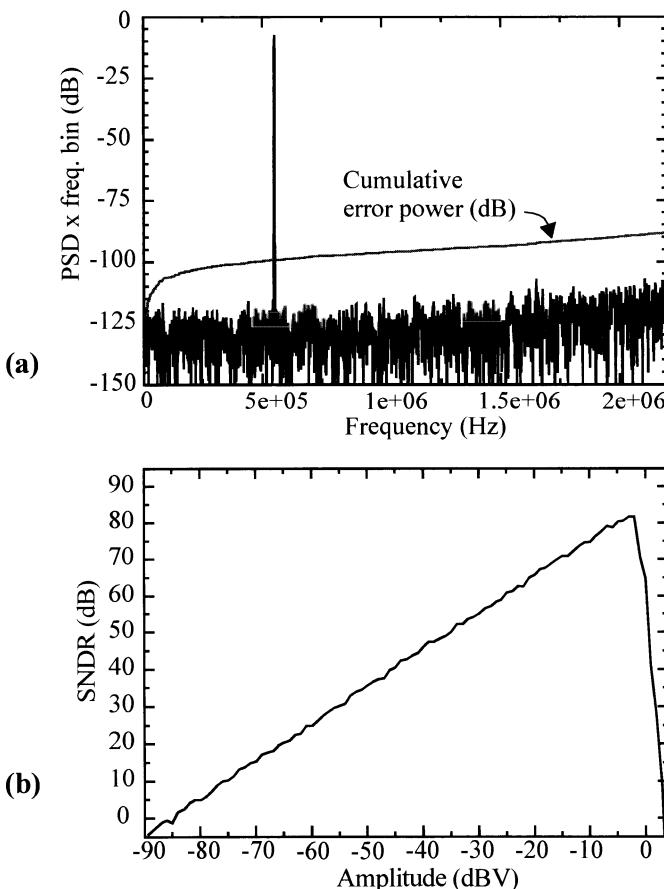


Figure 9.15: Simulation results: (a) output spectrum for -2dBV @550-kHz input sinewave. (b) SNDR vs. input amplitude.

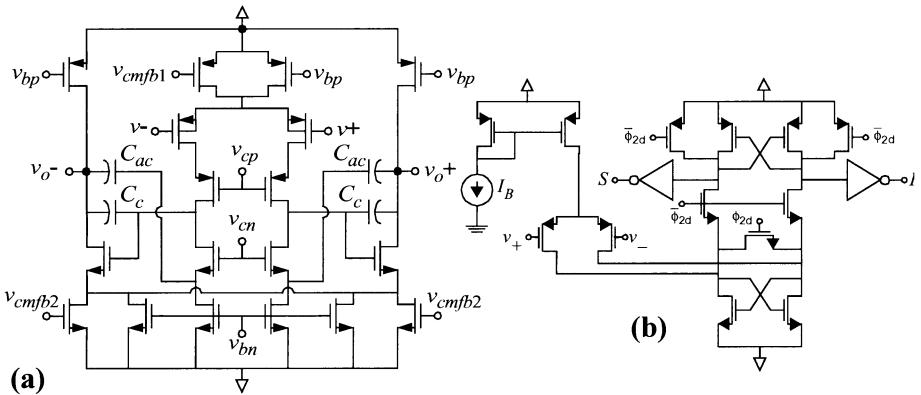


Figure 9.16: (a) Two-stage OTA. (b) Latched comparator.

shows the open-loop DC-gain of both OTAs as a function of the output voltage obtained from electrical simulations. A table look-up method implemented in ASIDES allows us to validate the actual DC-gain curve.

Comparators at the end of the modulator first and second stages require low resolution time (around 3ns) and hysteresis lower than 20mV. The regenerative latch including a pre-amplifying stage shown in Fig. 9.16(b) has been used [36].

The value of $R_{on} \sim 200\Omega$ can be obtained using CMOS switches with 1.8-V supply without clock-bootstrapping. Nevertheless, in low-voltage technologies the switch DC-characteristic is highly non-linear, causing a dynamic distortion, the more evident the larger the signal frequency [37]. The sampling process in the integrators is analyzed using electrical simulations for sine-wave or DMT signals. With the switch used, THD is always lower than -96dB for sinewaves (see Fig. 9.18) and has no practical influence on the MTPR [1] of DMT signals.

With the set of integrator weights used only 2×16 unitary capacitors are required. In the technology selected for this design example, the capacitors can be

Table 9.4: OTA simulation results

	OPA	OPB
Open-loop DC-gain	72.1dB	65.8dB
<i>GBW</i>	200MHz	148MHz
<i>PM</i>	60.3°	60°
<i>SR</i>	$442V/\mu s$	$385V/\mu s$
Differential <i>OS</i>	$\pm 1.35V$	$\pm 1.4V$
Input capacitance	0.2pF	0.1pF
Input equivalent noise	$2.54nV/\sqrt{Hz}$	$3.9nV/\sqrt{Hz}$
Power consumption	16mW	9mW

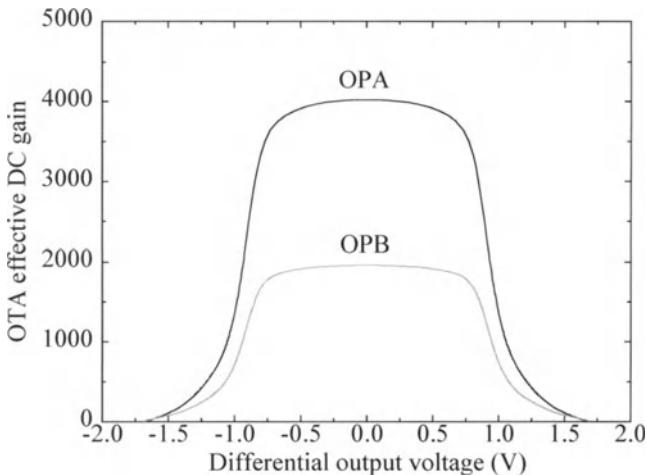


Figure 9.17: Illustrating the OTA DC-gain non-linearity.

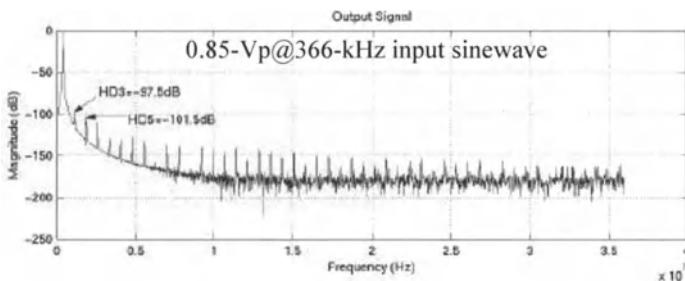


Figure 9.18: Harmonic distortion due to non-linear sampling.

implemented using metal-insulator-metal (M-i-M) structures, which are advantageous over other structures due to their good linearity and matching properties, as well as reduced coupling to substrate.

The 3-bit A/D/A converter consists of a simple fully-differential flash ADC and a resistive-ladder DAC [3][22]. The comparators in the ADC are of the same type as the ones at the end of the first and second stages.

9.4.4 Results

The complete modulator has been validated at the electrical level using HSPICE. The 8192-point FFT of the modulator output is shown in Fig. 9.19. The full-electrical simulation takes long time and considerable computational resources (4-day CPU time in this example) to obtain a reduced number of samples. For this reason, it is usually reserved for checking the connectivity of the

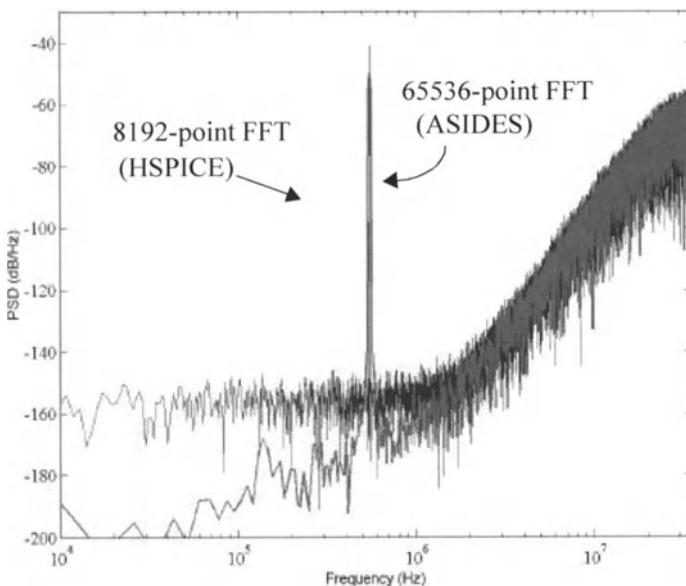


Figure 9.19: Simulated output spectrum

complete modulator schematic. For comparison purposes, the ASIDES simulation results (65536 samples in 4-sec. CPU time) has been added to Fig. 9.19. The complete modulator dissipates 67mW.

REFERENCES

- [1] H.J. Casier: "Requirements for Embedded Data Converters in an ADSL Communication System," in *Proc. of the IEEE ICECS'01*, vol. I, pp. 489-492, September 2001, Malta.
- [2] S. R. Norsworthy, R. Schereier and G. C. Temes (Editors): *Delta-Sigma Data Converters: Theory, Design and Simulation*, IEEE Press, 1996.
- [3] F. Medeiro, B. Pérez-Verdú and A. Rodríguez-Vázquez: "A 13-bit, 2.2-MS/s, 55-mW Multibit Cascade $\Sigma\Delta$ Modulator in CMOS 0.7- μ m Single-Poly Technology," *IEEE J. of Solid-State Circuits*, vol. 34, pp. 748-760, June 1999.
- [4] A.R. Feldman , B. E. Boser, and P. R. Gray: "A 13-Bit, 1.4-MS/s Sigma-Delta Modulator for RF Baseband Channel Applications," *IEEE J. of Solid-State Circuits*, vol. 33, pp. 1462-1469, October 1998.
- [5] T.L. Brooks, D. H. Robertson, D. F. Kelly, A. Del Muro, and S. W. Harston: "A Cascaded Sigma-Delta Pipeline A/D Converter with 1.25 MHz Signal Bandwidth and 89 dB SNR," *IEEE J. of Solid-State Circuits*, vol. 32, n. 12, pp. 1896-1906, December 1997.
- [6] I. Fujimori, L. Longo, A. Hairapetian, K. Seiyama, S. Kosic, J. Cao, and S.-L. Chan: "A 90-dB SNR 2.5-MHz Output-Rate ADC Using Cascaded Multibit Delta-

- Sigma Modulation at 8x Oversampling,” *IEEE J. of Solid-State Circuits*, vol. 35, pp. 1820-1828, December 2000.
- [7] K. Vleugels, S. Rabii, and B.A. Wooley: “A 2.5-V Sigma-Delta Modulator for Broadband Communications Applications,” *IEEE J. of Solid-State Circuits*, vol. 36, pp. 1887-1899, Dec. 2001.
 - [8] Y. Geerts, M.S.J. Steyaert, and W. Sansen: “A High-Performance Multibit $\Sigma\Delta$ CMOS ADC,” *IEEE J. of Solid-State Circuits*, vol. 35, pp. 1829-1840, December 2000.
 - [9] T.-H. Kuo, K.-D. Chen and H.-R. Yeng: “A Wideband CMOS Sigma-Delta Modulator with Incremental Data Weighted Averaging,” *IEEE J. of Solid-State Circuits*, vol. 37, pp. 2-10, January 2002.
 - [10] R. del Río, J. M. de la Rosa, F. Medeiro, B. Pérez-Verdú, and A. Rodríguez-Vázquez: “Top-down Design of a xDSL 14-bit 4-MS/s Sigma-Delta Modulator in Digital CMOS Technology,” in *Proc. of the Design, Automation and Test in Europe Conf.*, pp. 348 -351, March 2001.
 - [11] F. Medeiro, B. Pérez-Verdú, and A. Rodríguez-Vázquez: *Top-Down Design of High-Performance Sigma-Delta Modulators*, Kluwer Academic Publishers, 1999.
 - [12] F. Op'T Eynde and W. Sansen: *Analog Interfaces for Digital Signal Processing Systems*, Kluwer Academic Publishers, 1993.
 - [13] W.L. Lee and C.G. Sodini: “A Topology for Higher Order Interpolative Coders,” in *Proc. of IEEE International Symposium on Circuits and Systems*, pp. 459-462, 1987.
 - [14] S. M. Moussavi and B. H. Leung: “High-Order Single-Stage Single-Bit Oversampling A/D Converter Stabilized with Local Feedback Loops,” *IEEE Transactions on Circuits and Systems*, vol. 41, pp.19-25, January 1994.
 - [15] T. Cataltepe et al.: “Digitally Corrected Multi-bit $\Sigma\Delta$ Data Converters,” in *Proc. of IEEE International Symposium on Circuits and Systems*, pp. 647-650, 1989.
 - [16] M. Sarhang-Nejad and G.C. Temes: “A High-Resolution $\Sigma\Delta$ ADC with Digital Correction and Relaxed Amplifiers Requirements,” *IEEE J. of Solid-State Circuits*, vol. 28, pp. 648-660, June 1993.
 - [17] F. Chen and B. H. Leung: “A High resolution Multibit Sigma-Delta Modulator with Individual Level Averaging,” *IEEE J. of Solid-State Circuits*, vol. 30, pp. 453-460, April 1995.
 - [18] R. T. Baird and T. S. Fiez: “A Low Oversampling Ratio 14-b 500-kHz $\Delta\Sigma$ ADC with a Self-Calibrated Multibit DAC,” *IEEE Journal of Solid-State Circuits*, vol. 31, pp. 312-320, March 1996.
 - [19] O. Nys and R. Henderson: “A Monolithic 19bit 800Hz Low-Power Multi-bit Sigma Delta CMOS ADC using Data Weighted Averaging,” in *Proc. of European Solid-State Circuits Conference*, pp. 252-255, 1996.
 - [20] F. Chen and B. Leung: “A Multi-Bit $\Sigma\Delta$ DAC with Dynamic Element Matching Techniques,” in *Proc. of IEEE Custom Integrated Circuits Conference*, pp. 16.2.1-16.2.4, May 1992.
 - [21] Y. Matsuya et al.: “A 16-bit Oversampling A-to-D Conversion Technology Using

- Triple-Integration Noise Shaping," *IEEE J. of Solid-State Circuits*, vol. 22, pp. 921-929, December 1987.
- [22] B. Brandt and B. A. Wooley: "A 50-MHz Multibit $\Sigma\Delta$ Modulator for 12-b 2-MHz A/D Conversion," *IEEE J. of Solid-State Circuits*, vol. 26, pp. 1746-1756, December 1991.
 - [23] N. Tan and S. Eriksson: "Fourth-Order Two-Stage Delta-Sigma Modulator Using Both 1 Bit and Multibit Quantizers," *Electronics Letters*, vol. 29, pp. 937-938, May 1993.
 - [24] F. Medeiro, B. Pérez-Verdú, J.M. de la Rosa and A. Rodríguez-Vázquez: "Multibit Cascade $\Sigma\Delta$ Modulator for High-Speed A/D Conversion with Reduced Sensitivity to DAC Errors," *Electronic Letters*, vol. 34, No. 5, pp. 422-424, March 1998.
 - [25] V.F. Dias and V. Liberali: "Cascade Pseudomultibit Noise Shaping Modulators," *IEE Proceedings-G*, vol. 140, No. 4, pp. 237-246, August 1993.
 - [26] J.C Morizio et al.: "14-bit 2.2-MS/s Sigma-Delta ADC's," *IEEE J. of Solid-State Circuits*, vol. 35, pp. 968-976, July 2000.
 - [27] J.C.H. Lin: "TSMC 0.25 μ m Mixed-Signal 1P5M+ MIM Salicide 2.5V/5.0V Design Guideline", Taiwan Semiconductors Manufacturing Co.
 - [28] A. Marques, V. Peluso, M. Steyaert, and W. Sansen: "Analysis of the Trade-off between Bandwidth, Resolution and Power in Delta-Sigma Analog to Digital Converters," in *Proc. of IEEE ICECS*, vol. 2, pp. 153-156, Lisbon, September 1998.
 - [29] B. Razavi: *Design of Analog CMOS Integrated Circuits*, McGraw-Hill, 2001.
 - [30] F. Maloberti: *Analog Design for CMOS VLSI Systems*, Kluwer Academic Publishers, 2001.
 - [31] Y. Tsividis, *Mixed Analog-Digital VLSI Devices and Technology*, McG.-Hill, 1996.
 - [32] Y. Taur and T.H. Ning: *Fundamentals of Modern VLSI Devices*, Cambridge University Press, 1998.
 - [33] R. del Rio, F. Medeiro, B. Pérez-Verdú, and A. Rodríguez-Vázquez: "Reliable Analysis of Settling Errors in SC Integrators – Application to $\Sigma\Delta$ Modulators," *Electronics Letters*, vol. 36, pp. 503-504, March 2000.
 - [34] F. Goodenough: "Analog Technologies of all Varieties Dominate ISSCC," *Electronic Design*, vol. 44, pp. 96-111, February 1996.
 - [35] A. M. Marques, V. Peluso, M.S.J. Steyaert, and W. Sansen: "A 15-b Resolution 2-MHz Nyquist Rate $\Delta\Sigma$ ADC in a 1- μ m CMOS Technology," *IEEE J. of Solid-State Circuits*, vol. 33, pp. 1065-1075, July 1998.
 - [36] G.M. Yin, F. Op't Eynde, and W. Sansen, "A High-Speed CMOS Comparator with 8-b Resolution," *IEEE J. of Solid-State Circuits*, vol. 27, pp. 208-211, Feb. 1992.
 - [37] W. Yu, S. Sen, and B.H. Leung, "Distortion Analysis of MOS Track-and-Hold Sampling Mixers Using Time-Varying Volterra Series," *IEEE Trans. on Circuits and Systems II*, vol. 46, pp. 101-113, Feb. 1999.

Chapter 10

Continuous-Time Sigma-Delta for IF

F. Gerfers, Y. Manoli, M. Ortmanns

10.1 Introduction

Deep submicron processing, the reduction of the supply voltage as well as the increasing packaging density and the overall low power requirements of integrated circuits and systems raise the demand for novel circuit and design techniques. An important building block of such a mixed-signal system is an analog-to-digital converter (ADC). A multitude of different ADC architectures were introduced over the past decades [1] [2] [3]. The field of application of a particular ADC depends on the achievable conversion speed, accuracy, susceptibility to circuit imperfections, power requirements etc..

Flash structures provide the highest conversion speed in the range of hundreds of megahertz up to gigahertz [4], but at the expanse of high power and die area requirements. Consequently, efficient realizations are in general limited to resolutions of less than 8bit. Pipeline and oversampling topologies combine high speed with 12 to 16bit accuracy, while oversampled converters show some important advantages:

- They only need relatively low precision analog components. This is especially important for low voltage, deep-submicron technologies.
- No need of a precision sample-and-hold circuit.

The most popular oversampling ADC architecture is based on $\Sigma\Delta$ modulation. The overall design of such an ADC can be performed in the discrete-time (DT) or continuous-time (CT) domain. Both implementations show different advantages. Beside the significantly relaxed specifications for the anti-aliasing filter, which can be a very power- and area consuming building block, continuous-time modulators show a higher power-efficiency. This is because the gain-bandwidth product (GBW) and slew rate requirements of the used amplifiers are much lower compared to their discrete-time counterparts. Furthermore, the sampling operation takes place inside the modulator loop. As a result both sampling errors and out-of-band signals which alias into the passband are greatly suppressed by the loop gain in the passband. An additional benefit are the quiet virtual ground nodes, due to the reduced switching within the modulator.

On the other hand, CT modulators show a higher sensitivity to extra-loop delay and clock

jitter. The focus of this chapter are the different design issues of a continuous-time $\Sigma\Delta$ modulator. Sec. 10.2 illustrates the equivalence between a DT and CT modulator, considering the modified Z-transform, the principle of equivalent loop filters and the extension for cascaded architectures. The following section discusses the influence of circuit non-idealities with respect to the overall power consumption and their effect on the entire $\Sigma\Delta$ modulator. A detailed description of available error and mismatch cancellation and compensation techniques is shown in Sec. 10.4. Finally, a low power design strategy and the corresponding implementation of a low voltage low power CT $\Sigma\Delta$ modulator is presented.

10.2 DT/CT Modulator Equivalence

A $\Sigma\Delta$ modulator usually consists of a filter, an internal quantizer and a feedback DAC. Basically the quantization noise entering at the quantizer is suppressed by oversampling and additionally shaped to higher frequencies by a high-pass filter. Thereby this filter can be realized either in the discrete or the continuous time domain as described in [5].

Fig. 10.1 shows a typical block diagram of a CT and a DT modulator. There $\hat{H}(s)$ and $H(z)$ are the continuous and discrete time filters, respectively, and $\hat{R}_D(s)$ is the transfer function of the continuous time digital to analog converter. The main differences are:

- that in CT modulators switching occurs only at the quantizer.
- the filters are generally implemented in switched capacitor (SC) technique for the discrete time or with a CT integrator in the case of a continuous time modulator.
- the DT modulator needs an anti-aliasing filter in front of the system, which has to be very steep for low oversampling ratios.
- the DAC of the CT modulator forms the feedback signal with regard to time.

Now, a good point to start with is by explaining how to find an equivalent CT architecture for a given DT one, which shows the same noise transfer function. This can be done using one of the following ways.

10.2.1 Modified Z-Transformation

The overall behavior of a CT modulator loop is nonetheless discrete time due to the fact that the loop is sampled in time by the clocked quantizer (see Fig. 10.1). Furthermore, the design procedure of a CT $\Sigma\Delta$ modulator loop filter employs in general a DT loop filter. Thus, design and simulation of the ideal CT $\Sigma\Delta$ modulator can be done in discrete time domain [6] [7] to speed-up the overall design procedure. Subsequently, the originated DT open loop filter function $H(z)$ (Fig. 10.1(a)) can be replaced by a CT equivalent $\hat{H}(s)$ (Fig. 10.1(b)) with respect to the DAC feedback impulse response (Fig. 10.2). The two modulators are equivalent on condition that *the quantizer input voltages are the same at the sampling instances* [5] [8].

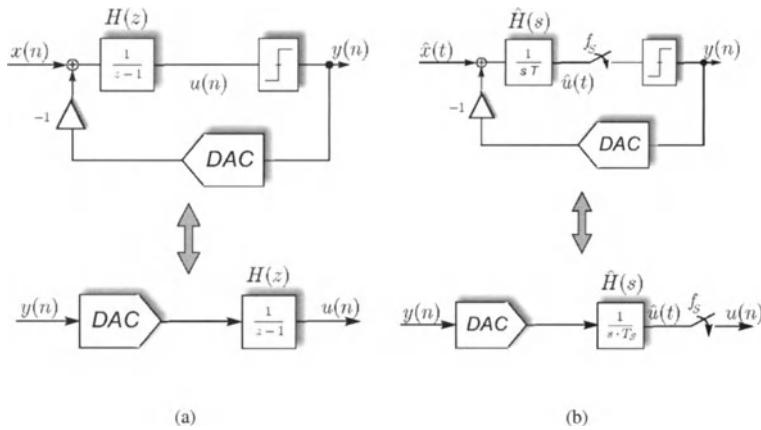


Figure 10.1: $\Sigma\Delta$ modulator block diagram and the corresponding open loop representation. (a) and (b) show the DT $\Sigma\Delta$ modulator and the CT equivalent.

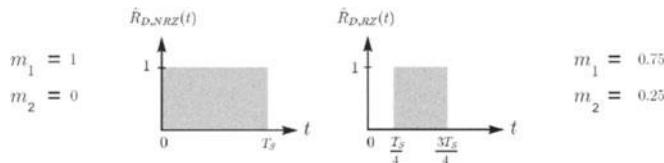


Figure 10.2: DAC feedback impulse responses $\hat{R}_D(t)$ in case of a NRZ-DAC (delay $\tau_d = 0$, pulse width $\tau = T_S$) and RZ-DAC ($\tau_d = 1/4T_S$, $\tau = 1/2T_S$) respectively.

The coefficients k_i of the CT modulator are calculated as a function of the DT integrator coefficients a_i by using for instance the modified Z-transform [6], [9], [10] and [11]. In the special case of a third order modulator (Fig. 10.3) the loop filter results in:

$$\mathcal{Z}(\hat{H}(s)) = \mathcal{Z}_{m1}(\hat{H}(s)) - \mathcal{Z}_{m2}(\hat{H}(s)) \quad (10.1)$$

$$\mathcal{Z}(\hat{H}(s)) = \mathcal{Z}_{m1}\left(\frac{-k_1}{TSs^2} + \frac{-k_2}{T^2s^3} + \frac{-k_3}{T^3s^4}\right) - \mathcal{Z}_{m2}\left(\frac{-k_1}{TSs^2} + \frac{-k_2}{T^2s^3} + \frac{-k_3}{T^3s^4}\right) \quad (10.2)$$

with $m_1 = 1 - \frac{\tau_d}{T}$ and $m_2 = 1 - \frac{\tau_d}{T} - \frac{\tau}{T}$. Using Tab. 10.1 and the NRZ-DAC feedback scheme ($m_1 = 1$, $m_2 = 0$) as well as the RZ-DAC feedback scheme ($m_1 = 3/4$, $m_2 = 1/4$), the resulting CT coefficients k_i are shown in Tab. 10.2, where a_1 , a_2 and a_3 are the

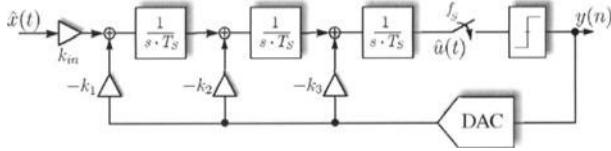


Figure 10.3: A general structure of a third-order single-bit CT $\Sigma\Delta$ modulator topology.

Table 10.1 Modified Z-transform for the respective loop filters.

$Z_m(\hat{H}(s))$	
$Z_m(1/s^2)$	$\frac{mT_S}{z-1} + \frac{T_S}{(z-1)^2}$
$Z_m(1/s^3)$	$\frac{T_S^2}{2} \left[\frac{m^2}{z-1} + \frac{2m+1}{(z-1)^2} + \frac{2}{(z-1)^3} \right]$
$Z_m(1/s^4)$	$\frac{T_S^3}{6} \left[\frac{m^3}{z-1} + \frac{3m^2+3m+1}{(z-1)^2} + \frac{6m+6}{(z-1)^3} + \frac{6}{(z-1)^4} \right]$

Table 10.2 CT loop filter coefficients for a NRZ and RZ-DAC.

	NRZ-DAC	RZ-DAC
k_1	$a_1 a_2 a_3$	$2a_1 a_2 a_3$
k_2	$a_1 a_2 a_3 + a_2 a_3$	$2a_1 a_2 a_3 + 2a_2 a_3$
k_3	$\frac{1}{3} a_1 a_2 a_3 + \frac{1}{2} a_2 a_3 + a_3$	$\frac{35}{48} a_1 a_2 a_3 + a_2 a_3 + 2a_3$

scaling coefficients of the DT modulator. The above example shows the simple handling of the *modified Z-transformation* on the one hand, but possesses indeed one main drawback on the other hand:

the *modified Z-transformation* does not accurate model the effect of excess loop delay, due to the fact that the *modified Z-transformation* assumes the delay happens at the output of the loop filter [8].

10.2.2 Equivalent loopfilters principle: the impulse-invariant transformation

In [7] a CT/DT modulator equivalence was proposed based on the following idea. A CT modulator as well as a DT modulator employs a clocked internal quantizer which makes both architectures discrete time systems. The only difference is the loop filter (LF), which is defined here as the path from the quantizer output to its input, and the signal representation in it. So finding a CT modulator equivalent to a DT architecture means designing a CT

loop filter with the same output as the DT counterpart at the quantizer sampling instants. This can be shown easily with Fig. 10.1. The conformity can be accomplished if the impulse responses of the loop filters from the quantizer outputs to their inputs are equal at the sampling instants.

$$Z^{-1}\{H(z)\} = L^{-1}\{\hat{R}_D(s)\hat{H}(s)\}|_{t=nT_S} \quad (10.3)$$

where $H(z)$ is the DT, $\hat{H}(s)$ the CT filter and $\hat{R}_D(s)$ the DAC transfer function respectively. From this relationship s-domain equivalences for z-domain LF poles can be determined, which has been done in [7] for a rectangular feedback DAC, i.e. RZ or NRZ, and has been extended for sloping feedback pulse forms in [12].

For the DT/CT transformation one has to convert the given DT loop filter into partial fractions resulting in a sum of poles $\frac{1}{(z-1)^i}$, $i = 1..n$, with n being the order of the DT loop filter. These poles can be transformed with the table given in [7] to CT poles, i.e. a CT loop filter. In [7] an example is shown for a 2nd order modulator, and a table for the transformation up to the order of 3 can be found there.

10.2.3 Extension for Multi Loop Modulators

The conversion of cascaded DT modulators to CT counterparts suffers from the fact that one stage feeds into another which requires the consideration of connecting loop filters. As mentioned in the previous section a $\Sigma\Delta$ -loop filter can be considered as a filter with the quantizer output as the input and the quantizer input as the output of the filter. This is straight forward for single loop modulators but can be easily extended to multi loop ones. *In this case not only each modulator stage results in a loop filter but also the paths from the outputs of previous stages to the quantizer inputs of following stages define further loop filters* [13].

The "Second-Order-First-Order (SOFO)" modulator is a common architecture to use for cascaded CT designs due to its reduced sensitivity to circuit nonidealities compared with other topologies as has been shown in [14]. Therefore the complete synthesis of this structure is shown here: The beginning point is to analyze the DT SOFO modulator proposed in [15] and all of its loop filters. The first one, $LF_1 = \frac{q_1}{y_1}$, is the loop filter of the first stage. The second one is defined the same way as $LF_2 = \frac{q_2}{y_2}$. Additionally the ‘connecting’ loop filter LF_3 from stage 1 to stage 2 has to be considered. This is the filter from the output of quantizer 1 (y_1) through the filters of the 1st and the 2nd stage to the

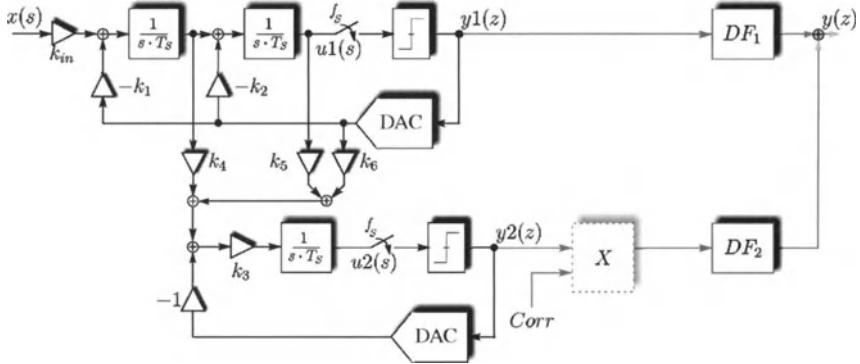


Figure 10.4: Block diagram of a third-order cascaded SOFO $\Sigma\Delta$ modulator topology ($M = 3, n_{loop} = 2$).

input of quantizer 2 (q_2). From [15] follows:

$$\begin{aligned} LF_1 &= -\frac{a_1 a_2}{(z-1)^2} - \frac{a_2}{z-1} \\ LF_2 &= -\frac{a_3}{z-1} \\ LF_3 &= \left(\frac{LF_1}{a_1 a_2} - b_1\right) \cdot c_1 \cdot \frac{a_3}{z-1} \\ &= -\frac{c_1 a_3}{(z-1)^3} - \frac{c_1 a_3}{a_1(z-1)^2} - \frac{c_1 a_3 b_1}{z-1} \end{aligned} \quad (10.4)$$

where a_1, a_2 and a_3 are the DT integrator scaling coefficients while b_1 and c_1 are the scaling coefficients between the first and the second stage.

These DT loop filters can be converted to their CT counterparts using the table given in [7] to get:

$$\begin{aligned} LFCT_1 &= -\frac{a_1 a_2}{s^2} - \frac{1}{2} \frac{a_2(1-a_1)}{s} \\ LFCT_2 &= -\frac{a_3}{s} \\ LFCT_3 &= -\frac{c_1 a_3}{s^3} - \frac{c_1 a_3 (\frac{1}{a_1} - 1)}{s^2} \\ &\quad - \frac{c_1 a_3 (b_1 + \frac{1}{3} - \frac{1}{2a_1})}{s} \end{aligned} \quad (10.5)$$

Now the resulting CT loop filters in Eq. 10.5 have to be realized by a CT modulator topology. As can be seen the loop filters of the 1st and the 2nd stage can be easily implemented

Table 10.3 Coefficients k_i for a CT SOFO modulator in Fig.10.4 in terms of the DT coefficients a_i , c_1 , b_1 .

k_1	$a_1 a_2$	k_2	$-\frac{1}{2} a_1 a_2 + a_2$
k_3	a_3	k_4	$-\frac{1}{2} \frac{c_1}{a_1 a_2}$
k_5	$\frac{c_1}{a_1 a_2}$	k_6	$-\frac{1}{6} \frac{c_1 (2a_1 + 6b_1 a_1 - 3)}{a_1}$

by CT modulators of orders 2 and 1 respectively. But if the same architecture is chosen in CT as in [15] for DT the equivalence can not be reached for the connecting loop-filter *LFCT*. To get a solvable set of equations the architecture has to be extended. Therefore the following statement can be made here which holds for all cascaded CT topologies we examined and derived with the presented strategy:

To realize the connecting loop filter from previous to following stages in cascaded CT $\Sigma\Delta$ modulators it is necessary to provide every integrator input of the later stages with every (weighted) CT state variable and the feedback DAC output of all previous stages.

For the CT SOFO modulator the resulting architecture is shown in Fig. 10.4. The corresponding weights k_i in terms of the DT modulator coefficients can then be calculated by comparison of the coefficients of the loop filter poles realized by the CT architecture shown in Fig. 10.4 and their DT/CT transformed counterparts in Eq. 10.5. The results for k_i are shown in Tab. 10.3. These and all following simulations are given for a NRZ DAC feedback pulse.

To show the effectiveness of this strategy synthesizing cascaded CT $\Sigma\Delta$ modulators the Signal-To-Noise(+Distortion)-Ratio (SNDR) of the CT SOFO modulator just developed will be shown. The simulations were performed using MATLAB and with specifications as follows: Audio $f_b = 25kHz$, $f_s = 1.6MHz \rightarrow OSR = 32$ and a 1 bit internal quantizer with output $\pm 1 \rightarrow$ quantizer step-size Δ_1 and $\Delta_2 = 2$ in the respective stage. The DT coefficients were chosen according to [15]. With these specs the in-band noise can be calculated to:

$$IBN = \frac{\pi^6}{7} \frac{\Delta^2}{12} \frac{1}{c_1^2 OSR^7} \approx -82,73dB \quad (10.6)$$

Fig. 10.5 compares the DT SOFO modulator of [15] with the derived CT one based on the achievable SNDR together with the ideal SNR based on Eq. 10.6. It is easily seen that there is nearly no difference between the DT and the CT SOFO modulator regarding the SNDR. Previously published methods for converting DT to CT cascaded modulators [16] have shown an acceptable matching of the two SNR as well, but no other transformation has shown this matching of the SNDR.

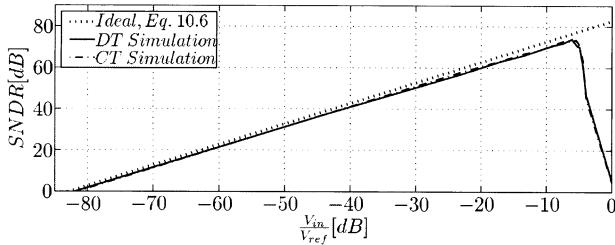


Figure 10.5: SNDR of an ideal, DT and CT SOFO topology

10.3 Influence of Circuit Non-Idealities

As specified in Section 10.2 a CT $\Sigma\Delta$ modulator consists of three building blocks. A CT loop filter, a clocked quantizer and a feedback DAC. Due to variations during the IC manufacturing process as well as through circuit imperfections their behavior deviates from the ideal behavior.

In this section the influence of a number of circuit non-idealities is analytically determined and verified by behavioral simulations in order to minimize their effect on the entire modulator performance on the one hand and to optimize the overall power consumption of each block on the other hand. In particular the effect of finite open-loop gain of the used amplifiers, non-dominant integrator poles, RC -mismatch which is equivalent to coefficient mismatch (Δk_i) as well as timing errors like excess loop-delay τ_d and clock jitter σ_t^2 will be discussed. Furthermore, the above specified circuit non-idealities can be divided into two categories [17]:

- a) those which alter the poles and zeros of the signal transfer function (STF) and noise transfer function (NTF): Finite open-loop gain, non-dominant integrator poles and RC -mismatch.
- b) and those which can be modelled as an error at the integrator input like circuit noise.

10.3.1 Integrator Non-Idealities

The first integrator is one of the most critical and power consuming circuits in a $\Sigma\Delta$ modulator, so its performance parameters should be considered in detail. Therefore, a substantial amount of power can be saved by a proper circuit design.

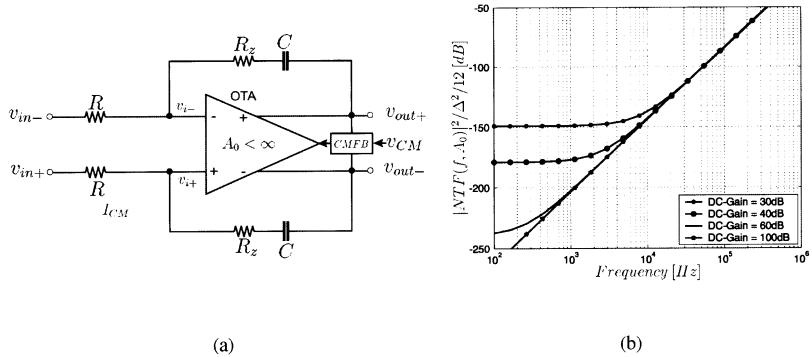


Figure 10.6: (a) Simplified schematic of a fully-differential integrator with finite open loop gain A_0 and excess phase cancellation (R_Z). (b) Influence of finite DC gain on the NTF of a third order CT $\Sigma\Delta$ modulator.

10.3.1.1 Leaky Integrator

In presence of finite amplifier gain A_0 one can express the CT integrator transfer functions $ITF_{\mathcal{E}_{A_0}}$ (see Fig. 10.6(a)) as:

$$ITF_{\mathcal{E}_{A_0}} = \frac{f_s A_0}{s(1 + A_0) + f_s} = \frac{\alpha f_s}{s + \gamma} \quad (10.7)$$

$$\alpha = \frac{A_0}{1 + A_0}, \quad \gamma = \frac{f_s}{1 + A_0} \quad (10.8)$$

where A_0 is the DC gain of the operational amplifier, $f_s = \frac{1}{RC}$ is the sampling frequency, while α and γ represent the integrator gain error and the pole displacement, respectively. Calculating now the total in-band quantization noise power (IBN) at the output of a third order ($M = 3$) single-loop modulator¹ one obtains [18]:

$$IBN(A_0) = \frac{\Delta^2}{12k_1^2 k_q^2} \cdot \left[\frac{\pi^6}{7OSR^7} + \frac{3\pi^4}{5OSR^5 A_0^2} + \frac{\pi^2}{OSR^3 A_0^4} + \frac{1}{OSR^1 A_0^6} \right] \quad (10.9)$$

$$IBN(A_0) \approx \frac{\Delta^2}{12k_1^2 k_q^2} \cdot \left[\frac{\pi^6}{7OSR^7} + \frac{3\pi^4}{5OSR^5 A_0^2} \right] \quad (10.10)$$

¹The nonlinear quantizer was modelled by a white-noise source (linear approximation).

and in general for an arbitrary order single-loop modulator [17]:

$$\text{IBN}(A_0) = \frac{\Delta^2}{12k_1^2 k_q^2} \cdot \left[\frac{1}{A_0^{2M} OSR} + \sum_{m=1}^M \left(\frac{\pi^{2m} M(M-1)\dots(M-m+1)}{(2m+1) OSR^{2m+1} A_0^{2(M-m)} m!} \right) \right]. \quad (10.11)$$

with k_q as effective quantizer gain [15]. The dominant mismatch term is in general of the order OSR^{2L-1} (Eqs. 10.9-10.10). The effect of finite amplifier gain on the NTF is shown in Fig. 10.6(b). It is obvious, that finite DC gain cause a shift of the zeros of the NTF away from DC. The resulting baseband noise is limited by the factor $\gamma \approx f_S/A_0 = 1/(A_0 RC)$ [19]. In order to minimize the additional noise caused by integrator leakage one can compute the critical DC gain value where the excess parasitic noise degrades the signal-to-noise ratio by 1dB. For the 3rd-order modulator this critical gain is equal to:

$$A_{0_{1dB}} \approx \sqrt{\frac{21}{5}} \cdot \frac{OSR}{\pi}. \quad (10.12)$$

Eq. 10.12 shows the minimum gain if only leakage is considered [18]. Taking also distortion into account the required DC gain has to be much higher (see Sec. 10.3.1.4). The requirements on the DC gain of single-loop topologies are still relaxed in contrast to multi-loop architectures [14]. For the SOFO structure the required gain A_0 has to be proportional to OSR^2 .

10.3.1.2 Time Constant Mismatch

Another critical imperfection in CT modulators are variations in the RC-time constant of the filters arising from large tolerances of integrated resistors and capacitors.

In first order approximation the variation of the time constant can be modelled by a relative error ϵ_{RC} and the ITF becomes:

$$\text{ITF}_{\epsilon_{RC}} = \frac{1}{s RC} = \frac{\frac{f_S}{1+\epsilon_{RC}}}{s} \quad (10.13)$$

This is not of such an issue in single loop modulators. Here the IBN of a 3rd order modulator due to RC-tolerance results in:

$$\text{IBN} \approx \frac{\Delta^2 \pi^6}{12} \frac{1+6\epsilon_{RC}}{k_1^2 k_q^2 OSR^7} \quad (10.14)$$

where variations of about ±20% result in only 3dB increase of the IBN [18]. Simulation results are given in Fig. 10.7

Contrarily in cascaded modulators some part of lower order shaped quantization noise of the first stage will leak through the overall modulator output [20]. The in-band noise for a 3rd order SOFO modulator in Fig. 10.4 can be calculated to:

$$IBN_{\epsilon_{RC}} \approx \frac{\pi^6}{7} \frac{\Delta_2^2}{12} \frac{1}{c_1^2 OSR^7} + \frac{4\pi^4}{5} \frac{\Delta_1^2}{12} \frac{\epsilon_{RC1}^2}{OSR^5} \quad (10.15)$$

If 3rd order noise shaping is required, it can be seen that ϵ_{RC} of the 1st stage has to be in the order of OSR^{-1} , which cannot be reached for usual OSR values. This is even worse if a multi-bit quantizer is used in the 2nd stage, as can be seen in Fig. 10.8 for the SOFO modulator. Here the simulated and calculated IBN is plotted over ϵ_{RC} for different resolutions of the 2nd quantizer. It can be seen, that the benefit of using a multi-bit quantizer in the 2nd stage is diminished even for small variations.

Despite of this absolute time constant mismatch also the much smaller relative mismatch between resistors or capacitors have to be considered. But both simulations and calculations show that this mismatch error has almost no influence on single loop CT modulators, whereas the influence is negligible in cascaded modulators compared to the absolute RC-gain mismatch.

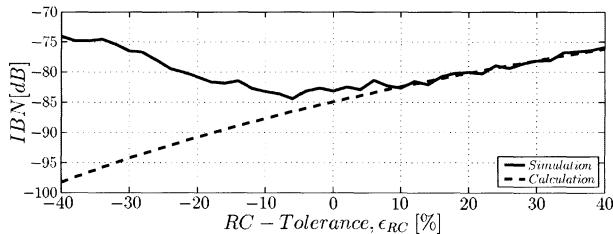


Figure 10.7: IBN of a CT 3rd order single loop modulator with RC tolerance

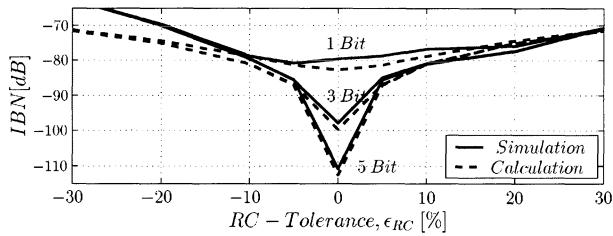


Figure 10.8: IBN of a CT SOFO with RC tolerance

10.3.1.3 Non-dominant Integrator Poles

Non-dominant integrator poles and their cause, finite gain bandwidth (GBW) of the amplifiers is more of an issue in DT than in CT circuitry as mentioned in [5] and [7]. Generally non-dominant integrator poles degrade the loop stability of the entire modulator by pushing the modulator poles closer to the imaginary axis [21]. Taking only the dominant pole ($p_{dom}(s)$) of the amplifier into account one can express the inband noise at the output of a 3rd order, single loop modulator as follows [18]:

$$IBN \approx \frac{\Delta^2}{12} \cdot \frac{\pi^6}{7 k_1^2 k_q^2 OSR^7} \cdot \left[1 + \frac{f_S}{GBW}\right]^6 \quad (10.16)$$

with $GBW \approx A_0 \cdot p_{dom}$

To limit the additional noise due to the non-dominant integrator pole to 3dB the minimal ratio of the amplifier GBW to the sampling frequency f_S is shown in Eq. 10.17:

$$c_{min} = \frac{GBW_{min}}{2\pi f_S} \approx \frac{1}{2\pi} \frac{1}{2^{\frac{1}{6}} - 1} \approx 1.3 \quad (10.17)$$

For cascaded modulators we have to investigate the influence more carefully [20]. Modelling the CT integrator using a single pole amplifier with high enough open loop gain A_0 as in Section 10.3.1.1, the ITF can be approximated to be:

$$ITF_{GBW} \approx \frac{f_S}{s} \cdot \frac{\frac{GBW}{GBW+f_S}}{\frac{s}{GBW+f_S} + 1} \quad (10.18)$$

So modelling a finite GBW error means first an integrator gain error $\frac{GBW}{GBW+f_S}$, which has the same influence as a variation ϵ_{RC} shown in Section 10.3.1.2, and secondly a non-dominant pole ($GBW + f_S$), whose influence has to be additionally modelled. Analytical results are extensive and not very didactically, therefore simulation results are given for the SOFO modulator here. Fig. 10.9 shows the IBN over the $c = \frac{GBW}{2\pi f_S}$ ratio, split into the two error sources, i.e. 2nd pole or gain error only, both for single-bit and multi-bit quantizer in the second stage. It is clearly seen that the pole influences the performance much less than the integrator gain error. Again the performance loss is worse when using second stage multi-bit quantization.

10.3.1.4 Other Integrator Non-Idealities

Beside the above presented integrator error mechanisms, there are still some further important error sources, Fig. 10.10. In general the DR and SNR of a $\Sigma\Delta$ modulator is

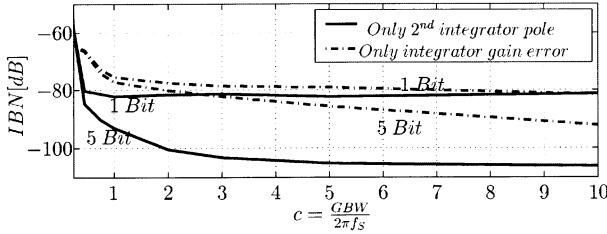


Figure 10.9: IBN of a CT SOFO with finite GBW

limited by circuit noise, here by $\bar{v}_{n,OTA}^2$, $\bar{v}_{n,R}^2$, \bar{v}_{n,R_Z}^2 . Consequently, the ADC resolution can be specified by means of calculating the total input referred noise. If required, chopper techniques [22] can be applied to reduce these effects in the passband, but at the expense of a sampled integrator. Furthermore, the voltage dependency of the amplifier DC

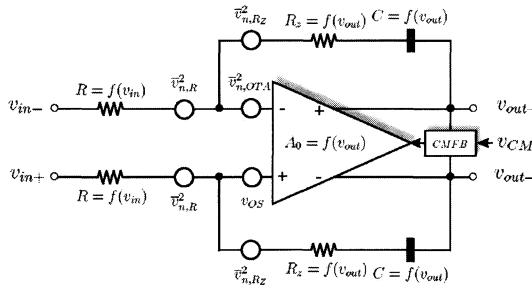


Figure 10.10: Simplified schematic of a fully-differential integrator with additional error and noise sources.

gain, integrator resistances and capacitors may introduce substantial distortion. Fig. 10.11 shows the simulated open loop gain of a single-stage folded-cascode amplifier [23]. The maximum gain of 8050 is obtained at the common mode output voltage and decreases rapidly as the output voltage approaches the saturation region. Estimating the nonlinearity one obtains a THD of less than $-99dB$ according to [24].

10.3.2 Temporal Non-Idealities

10.3.2.1 Excess Loop Delay

Excess loop delay τ_d is defined as the *time* between the quantizer decision and the DAC pulse, Fig. 10.12. In [7] and [8] it is shown that in presence of any delay the loop transfer

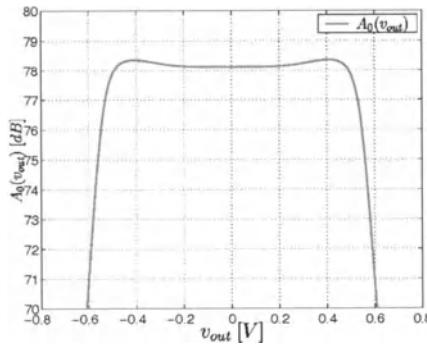


Figure 10.11: Simulated amplifier gain nonlinearity.

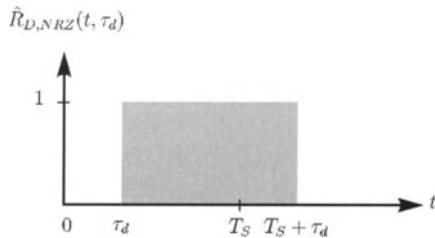


Figure 10.12: DAC feedback impulse response $\hat{R}_{DAC,NRZ}(t, \tau_d)$ under the influence of excess loop delay τ_d .

function, the in-band noise as well as the maximum stable signal amplitude will be affected. This will be evident, by estimating the influence of loop delay on the location of the poles and zeros of the loop filter. Considering excess loop delay τ_d the DT loop filter $H(z, \tau_d)$ of a general CT third order modulator (Fig.10.3) results in :

$$\begin{aligned}
 H(z, \tau_d) = & -\frac{1}{6} \frac{[(\tau_d - 1)^3 k_1 - 3(\tau_d - 1)^2 k_2 + 6(\tau_d - 1)k_3]z^3 + \dots}{[(-3\tau_d^3 + 6\tau_d^2 - 4)k_1 + 3(3\tau_d - 4)\tau_d k_2 + 6(-3\tau_d + 2)*k_3]z^2 + \dots} \\
 & \frac{[(3\tau_d^3 - 3\tau_d^2 - 3\tau_d - 1)k_1 + 3(-3\tau_d^2 + 2\tau_d + 1)k_2 + 6(3\tau_d - 11)k_3]z + \dots}{-\tau_d^3 k_1 + 3\tau_d^2 k_2 - 6\tau_d k_3} \\
 & \frac{}{z(z - 1)^3} \quad (10.19)
 \end{aligned}$$

For this, the impulse-invariant transformation was used, according to Sec. 10.2.2. The entire loop delay was modelled as a delayed DAC feedback pulse in the present sampling instant plus a feedback pulse of the duration τ_d at the next sample [8]. From Eq. 10.19 it is obvious, that the loop filter order is increased as well as the location of the zeros are shifted. Consequently, excess loop delay affects beside the overall modulator performance also the loop stability.

The effect of extra loop-delay on the NTF poles is shown in detail in Fig. 10.13(a). For this, the excess delay τ_d was increased from $0.01\%T_S$ to $90\%T_S$. Fig. 10.13(a) displays that a delay exceeding $78\%T_S$ results in an unstable modulator whereas the behavioral simulation (Fig. 10.13(b)) indicate that a delay of just $\leq 1\%T_S$ is tolerable without leading to any performance degradation.

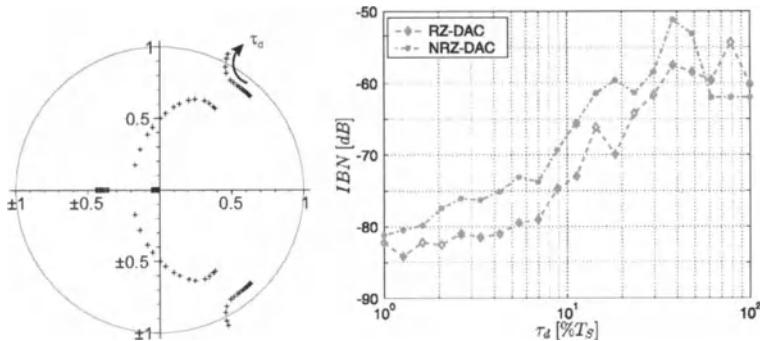


Figure 10.13: (a) Influence of extra loop delay τ_d on the poles of a third-order modulator with $k_1 = 0.05$, $k_2 = 0.3$ and $k_3 = 0.64$ and (b) the corresponding behavioral simulation result.

10.3.2.2 Clock Jitter

When considering high speed CT designs, one of the most critical parameters is clock jitter, i.e. statistical variations of the clock edges. The magnitude of these variations depends on the clock source used for the modulator, and if a standard digital clock (e.g. from a PLL) has to be used for economic reasons, jitter even in the hundred ps range has to be expected, limiting possible sampling frequencies to very low values [18] [25].

The timing errors due to clock jitter noise in the feedback loop increase the noise level in the signal band. For typical rectangular feedback pulse forms, e.g. RZ or NRZ, a statistical variation of the clock width means a variation of the pulse length. In Fig. 10.14 it is easily

seen that the corresponding error is proportional to the timing error. According to [21]

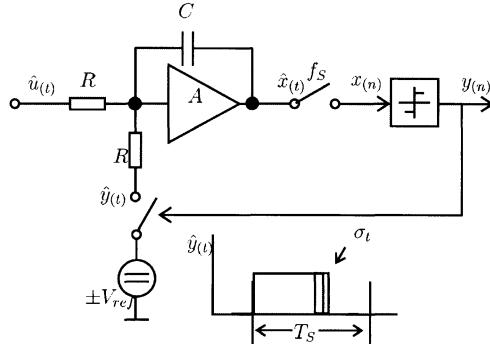


Figure 10.14: CT $\Sigma\Delta$ modulator with rectangular feedback DAC

and [26] the in-band jitter noise power can be approximated for a rectangular DAC pulse to:

$$IBN_{jitter} = \frac{\Delta^2 (\frac{\sigma_t}{T_S})^2}{OSR} \quad (10.20)$$

with σ_t^2 being the variance of the clock jitter. Eq. 10.20 is slightly varying between NRZ and RZ feedback DAC [26].

For a 3rd order modulator with $OSR = 48$ and $k_1 = 0.05$ a maximum clock jitter of $\sigma_t \approx 0.025\% T_S$ is tolerable for a 3dB loss in performance according to [18].

This degradation of the modulator performance due to the variation in width of the feedback pulse superposes all other jitter effects, i.e. due to pulse position jitter [27] or the whitening of the signal due to accumulated jitter [25]. A technique to reduce the sensitivity of CT modulators to clock jitter has been presented in [12] and will be shown in Sec. 10.4.2.

10.3.3 Other Non-Idealities

Beside the above specified error mechanisms, errors in the ADC and primarily in the DAC can affect the overall modulator resolution. While the non-ideal behavior of the ADC is greatly suppressed by the loop gain [28], errors in the DAC are directly input-referred [8] [3]. Thus, e.g. the linearity of a multibit DAC has to be equivalent to the overall modulator resolution [29]. However, there is no significant difference among a continuous time and discrete time $\Sigma\Delta$ modulator in this respect.

10.4 Error Compensation and Cancellation

Due to the fact that CT modulators suffer more from some non-idealities as their DT counterparts the demand arises to somehow compensate for this performance degradation. Even more urgent is the cancellation of those effects, which make the implementation of some CT modulators impossible. Therefore a reduction should be achieved in:

- the clock jitter sensitivity of CT modulators.
- the enormous performance degradation of cascaded CT modulators due to integrator gain errors.

10.4.1 Cancellation of Gain Mismatch in Cascaded Modulators

It has been shown [20] [30], that errors due to variations in the RC-time constant can be eliminated in the digital part of a cascaded modulator. When comparing analytically the outputs of both quantizers of a cascaded modulator as in Fig. 10.4 it can be shown, that lower order shaped noise due to RC-variations differs only by a factor. Inversely multiplying that factor to the output bit stream, shown as dotted lines in Fig. 10.4, cancels out the excess noise due to RC-variations.

In general the gain error cancellation means that any integrator gain error must be measured and eliminated by multiplying the output of the 2nd quantizer by the inverse of the squared error. "Corr" in Fig. 10.4 then becomes:

$$ITF_{g_{err}} = \frac{f_s}{s} \cdot g_{err}, \quad Corr = \frac{1}{g_{err}^2} \quad (10.21)$$

whereby the gain error may be caused by integrator time constant mismatch as in Eq. 10.13 or due to finite GBW as in Eq. 10.18.

Recalculating the IBN of Eq. 10.15 while using this error correction technique, one gets:

$$IBN_{\epsilon_{RCcorr}} = \frac{\pi^6}{7} \frac{\Delta_2^2}{12} \frac{(1 + 6\epsilon_{RC} + 15\epsilon_{RC}^2 + 20\epsilon_{RC}^3)}{c_1^2 OSR^7} \quad (10.22)$$

Most important is that all parts of the IBN in Eq. 10.22 are 3rd order noise shaped. Simulation results corresponding to RC variations as in Fig. 10.8, but employing this technique, have already been shown in [13].

Adopting the cancellation technique for the finite GBW induced gain errors the IBN can be expressed with the same equation as in Eq. 10.22, but with ϵ_{RC} replaced by $\frac{f_s}{GBW}$. So it can be clearly seen, that the technique is also very useful in decreasing the required GBW of the amplifiers in cascaded CT modulators [20].

10.4.1.1 Limitations of the Error Cancellation

The presented technique implies, that all integrators used in the modulator have the same gain and this gain can be exactly determined. In reality neither the matching nor the measurement will be perfect, which can be modelled by replacing the proposed values of "Corr" above by:

$$Corr = \frac{1}{(g_{err} \cdot (1 + \epsilon_{Corr}))^2} \quad (10.23)$$

what implies a correction error ϵ_{Corr} in the determined value of the integrator gain error g_{err} corresponding to the definition in Eq. 10.21.

A rough but realistic lower limit estimation for this detection error can be derived from relative matching of integrated RC-products. Taking about 0.5-1% relative matching for R and C, respectively, the product matches within about 1% - 2%.

Exact analytical results are too extensive to be shown, but a good estimation for the IBN is:

$$IBN_{\epsilon_{RC_{Errcorr}}} = IBN_{\epsilon_{RC_{corr}}} + \frac{4\pi^4}{5} \frac{\Delta_1^2}{12} \frac{(\epsilon_{Corr}^2)}{OSR^5} \quad (10.24)$$

In Eq. 10.24 again 2nd order shaped noise appears due to the correction error ϵ_{Corr} , but as apposed to Eq. 10.15 ϵ_{Corr} here is the relative matching error, which is much smaller than the RC-variation ϵ_{RC} itself. Simulation results are shown in Fig. 10.15. Here again the SOFO with gain error correction is simulated for single and multibit 2nd quantizer, first without detection error, secondly with $\epsilon_{Corr} = 1\%$ and thereunto the calculated results from Eq. 10.24. For the chosen detection error it is obvious that there is nearly no difference with or without correction error if a single bit quantizer is used in the 2nd stage, that for higher quantizer resolution in this stage the ideal IBN can't be reached. But a very good overall performance can be obtained for a reasonable value for the detection error. These results also suggest being careful in using higher order CT cascaded topologies like a 2-2 or 2-1-1 modulator [15]. Their lower (2nd and 3rd) order noise would remain in the signal band due to the correction error, thus lowering the IBN to levels already achieved by the SOFO modulator, a much smaller system.

The presented results can also be adopted if a gain error due to finite GBW is assumed.

10.4.2 CT Modulators with reduced Clock Jitter Sensitivity

10.4.2.1 Clock Jitter in DT Modulators

The relative insensitivity to clock jitter of DT $\Sigma\Delta$ modulators realized in switched capacitor (SC) technique is based on the sloping pulse form of the feedback. This slope

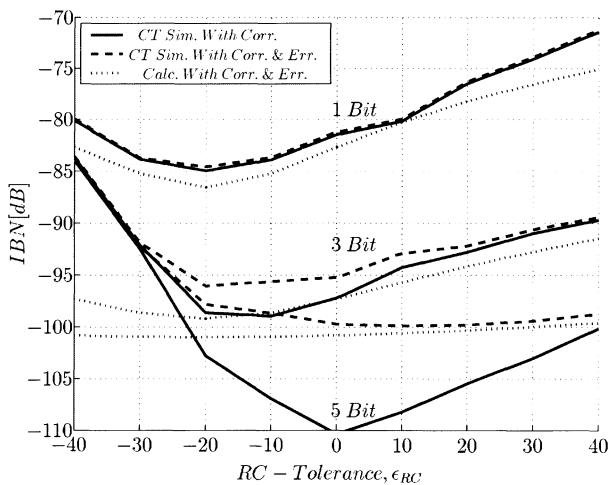


Figure 10.15: IBN of a SOFO with non-ideal correction

arises from the capacitor being discharged over a switch with very low 'on-resistance'. Therefore the slope is very steep, following the time constant:

(10.25)

A feedback signal of this form is shown in Fig. 10.16. Here it can be seen, that the charge error of the feedback signal due to clock jitter is much lower as it was in Fig. 10.14.

10.4.2.2 Using SC Technique for CT Modulators

In [12] a modified SC feedback is proposed to be used in CT modulators as well. Here the signal path is still switched only at the quantizer, and the CT modulator still implies

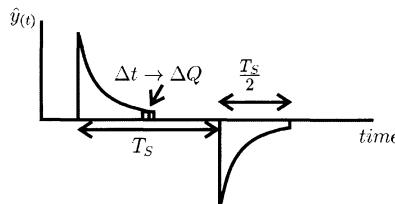


Figure 10.16: Capacitor discharging pulse form

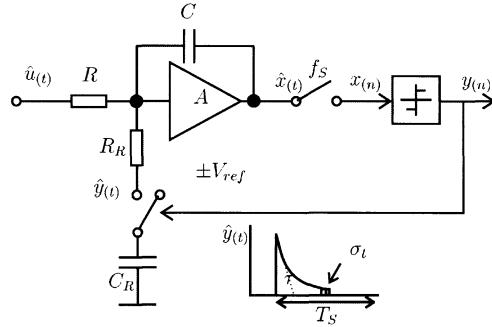


Figure 10.17: CT $\Sigma\Delta$ modulator with SCR feedback DAC, corresponding to Fig.10.14

the intrinsic anti-aliasing filter.

On the other hand the required bandwidth and the Slew-Rate (SR) of the filter integrators have to be much higher due to the very fast SC-signal pulse, canceling some of the advantages of the CT modulator.

To avoid this fast switching it is proposed to discharge the capacitor C_r not directly to the integrator input, but adding an additional resistor R_r in series and defining the resulting structure as *SCR Feedback*. Thus the discharge current of the capacitor is:

$$I_C = \frac{U_0}{R_r} \exp^{-\frac{t}{\tau}}, \quad \tau = R_r \cdot C_r \quad (10.26)$$

An exemplary CT modulator is shown in Fig.10.17, where $f_s = \frac{1}{RC}$.

10.4.2.3 Synthesis of the SCR Feedback

When designing a CT modulator to perform a given noise transfer function, the synthesis procedure can be done following the method of the *Equivalent Loopfilters* described in Section 10.2.2.

Here as well as in [7] the transformation procedure is shown and a table for equivalent loopfilter poles for a rectangular DAC pulse is given. In [12] an equivalent table is given for the presented SCR feedback, where the pulse form is assumed as in Fig. 10.16.

Following the procedure described in [7] but using the new table in [12] for the transformation, the coefficients for the 2nd order modulator of Fig. 10.18 implemented with the proposed SCR feedback can be exemplary derived for a feedback pulse position $\alpha = 0$

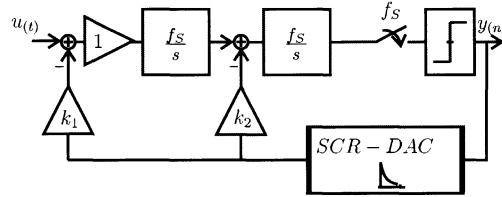


Figure 10.18: Block diagramm of a 2nd order CT modulator

and $\beta = 0.5$, i.e. in the first half of the clock phase:

$$k_1 = \frac{T_S}{\tau (1 - e^{-\frac{T_S}{2\tau}})} \quad (10.27)$$

$$k_2 = \frac{3T_S + 2\tau - 2 e^{\frac{T_S}{2\tau}} (T_S + \tau)}{2\tau (2 - e^{\frac{T_S}{2\tau}} - e^{-\frac{T_S}{2\tau}})} \quad (10.28)$$

With the given table also higher order and even cascaded modulator structures can be synthesized [13].

Fig. 10.19 shows a power spectrum density for the 2nd order modulator of Fig. 10.18 with $f_S = 40.32\text{MHz}$, baseband $f_B = 420\text{kHz}$, $f_{Signal} = 200\text{kHz}$, an OSR = 48 and for example $\tau = 0.2 T_S$. Additionally the integrated IBN is shown. The proper function of the synthesized modulator can be clearly seen.

10.4.2.4 Jitter Sensitivity of the SCR feedback

The lower sensitivity to clock jitter of the proposed technique is seen when comparing Fig. 10.14 to Fig. 10.17, and has been proven analytically in [31] similar to [21]. The

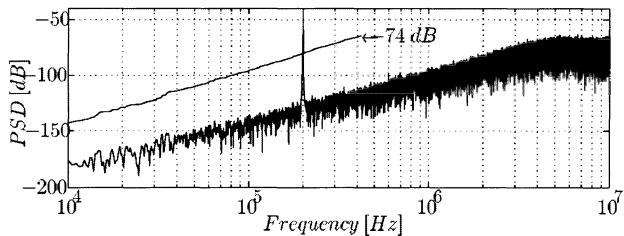


Figure 10.19: PSD of a 2nd order CT SCR modulator

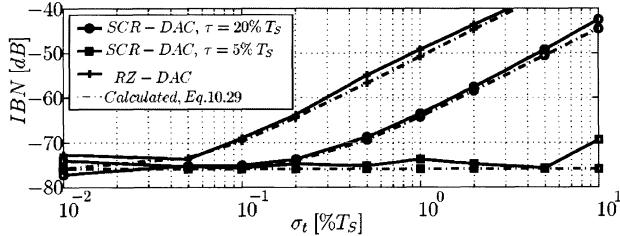


Figure 10.20: Inband noise due to clock jitter in a 2nd order modulator with RZ or SCR feedback

resulting IBN due to clock jitter is:

$$S_{\Delta t} = \frac{\sigma_j^2}{OSR} = \frac{\Delta^2}{4} \left(\frac{T_s}{\tau}\right)^2 \left(\frac{\sigma_t}{T_s}\right)^2 e^{-\frac{T_s}{\tau}} \frac{1}{OSR} \quad (10.29)$$

In Fig. 10.20 the presented technique is confirmed by simulation of the SCR feedback modulator (Fig. 10.18) with two different values of the feedback time constant, $\tau = 0.2 T_s$ and $\tau = 0.05 T_s$, compared with a modulator with rectangular RZ feedback DAC. Additionally the theoretical results of Eq. 10.20 and Eq. 10.29 are included. Clock jitter is given in percentage of the sampling time. Specifications are chosen as in the example of Fig. 10.19.

At a clock jitter variance of 1% T_s the RZ-DAC modulator suffers an IBN penalty of about 25dB, whereas the SCR-DAC modulator, depending on the settling time τ , has no or just a 10dB decline in IBN.

Finally in [32] an implementation of a 3rd order CT modulator with SCR feedback is presented and the usefulness of the technique is experimentally proven.

10.4.2.5 Constraints of the SCR feedback

It is obvious that there are some constraints concerning speed, when using the proposed technique. This is due to the pulse form of the capacitor discharge current which is sloping at the end, but heavily peaking at the beginning of the pulse requiring a higher slew rate (*SR*) of the amplifiers. But compared to a DT modulator the required SR is smaller due to the proposed resistor R_r . Fig. 10.21 shows the Matlab simulated IBN of the modulator in Fig. 10.18 as well as a modulator with RZ feedback with $f_s \approx 40MHz$, when using finite SR. It can be seen that the minimal required SR becomes slightly higher with lower τ and therefore with better jitter insensitivity.

An additional speed constraint is the higher gain bandwidth product required, when using the proposed SCR feedback, compared with usual CT modulators. This is shown in

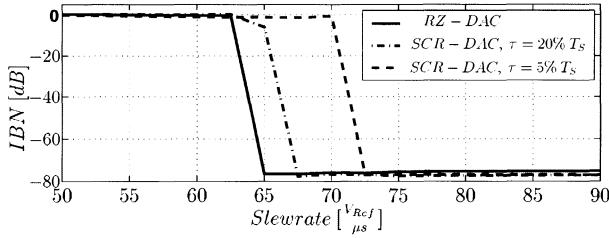


Figure 10.21: Influence of finite slew rate for a 2nd order modulator with SCR- and RZ-DAC

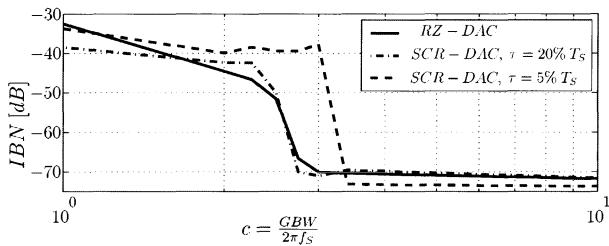


Figure 10.22: Influence of finite GBW

Fig. 10.22. Again it can be seen that amplifier speed demands become somewhat higher, when choosing smaller values for τ .

10.5 Implementation of Low Power CT $\Sigma\Delta$ Modulators

In this section, a design strategy for low-power continuous-time $\Sigma\Delta$ converters is proposed. This concept is based on the Figure of Merit (*FOM*) which takes the overall power consumption (P), the dynamic range (DR) and the signal bandwidth (f_B) into account to find the most power-efficient $\Sigma\Delta$ modulator implementation with respect to these design parameters (Sec. 10.5.1). This strategy considerably simplifies the design flow of $\Sigma\Delta$ converters and makes it more systematic. Finally, the theoretical results will be verified in Section 10.5.4 by experimental results.

10.5.1 Figure of Merit, FOM

The *FOM* estimation comprises a mixed bottom-up top-down approach [18], [33]. The required parameters like dynamic range, signal bandwidth, supply voltage etc. will

be taken from different abstraction levels. The used *FOM* is defined as:

$$FOM = \frac{P}{DR \cdot 2f_B} \quad (10.30)$$

The dynamic range calculations consider only quantization (N_Q) and thermal noise (N_{Th}):

$$DR \approx \frac{\frac{(2\hat{V}_{sw})^2}{2}}{N_Q + N_{Th}} = 3 \cdot 2^{2B-1} \quad (10.31)$$

$$N_Q \approx \frac{(2\hat{V}_{sw})^2}{12} \frac{\pi^{2M}}{(2M+1) OSR^{2M+1}} \quad (10.32)$$

$$N_{Th} = \frac{kT}{OSR \cdot C} \quad (10.33)$$

where M is the modulator order, \hat{V}_{sw} the peak input voltage, k the Boltzmann constant, T the temperature and C the entire integration capacitance. The overall power consumption P is made up of a static and dynamic portion of the analog part (P_{static} , $P_{dynamic}$) and a portion of the digital part ($P_{digital}$).

$$P \approx P_{static} + P_{dynamic} + P_{digital} \quad (10.34)$$

$$P \approx n_I I_B V_{DD} M + 2\hat{V}_{sw}^2 C f_s M + n_Q P_{inv} f_s \quad (10.35)$$

where n_I represents the number of current branches times the current I_B . We assume, for reasons of simplicity, that the current is the same in every branch. n_Q is the size of the quantizer and additional logic converted in CMOS gate-equivalent inverters each with a power consumption of P_{inv} . This way, the entire power consumption results to Eq. 10.35. Thus, with Eqs. 10.32 to 10.33 one can calculate the required capacitance C , and accordingly with Eq. 10.35 the *FOM*. The optimal $\Sigma\Delta$ modulator implementation with respect to a minimal power consumption can be determined by $\frac{\partial FOM}{\partial OSR} = 0$.

During the last 15 years a large number of $\Sigma\Delta$ modulator architectures have been proposed, for example [34] [35] [36] [37] [38]. The main objective of all these schemes is the reduction of the quantization noise within the baseband. In the following the most common architectures will be analysed.

10.5.1.1 Single-Loop Architectures

In Fig. 10.3 a single-loop third order $\Sigma\Delta$ modulator architecture is shown. The drawback of single-loop modulators with $M > 2$ is the tendency to instability. Hence, a proper selection of the scaling coefficients must be done. A more realistic *FOM* estimation takes also this scaling of the noise transfer function (NTF) into account. For the *FOM*

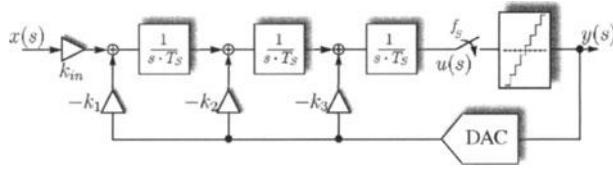


Figure 10.23: A general structure of a third-order single-bit CT $\Sigma\Delta$ modulator topology with a B_{int} bit quantizer.

computation a rule-of-thumb with an out-of-band gain ($\hat{H}_\infty = 2^M$) for the NTF of 1.5 was used.

$$N_Q^* \approx \frac{2^{M+1}}{3} N_Q \quad (10.36)$$

The scaling of the $\Sigma\Delta$ modulator implies that the resulting in-band noise (IBN) is much higher than that calculated with Eq. 10.32. Consequently, the oversampling ratio has to be increased. This results in a higher power drain and accordingly in a higher FOM . A more efficient way to stabilize high order $\Sigma\Delta$ modulators is to increase the number of levels of the internal quantizer at the expense of linearity problems in the DAC.

10.5.1.2 Multi-Bit Single-Loop Architectures

The use of a multi-bit ADC (Fig. 10.23) as the internal quantizer offers beside a better loop stability also the advantage that the in-band quantization noise is a factor of $(2^{B_{int}} - 1)^2$ lower in comparison with a simple two level comparator:

$$N_Q^{**} \approx \frac{N_Q}{(2^{B_{int}} - 1)^2} \quad (10.37)$$

where B_{int} indicates the number of bits of the internal ADC. Furthermore, the power consumption of the digital part is roughly increased by the resolution of the internal ADC. Therefore, the digital power dissipation expands to:

$$P_{digital}^* \approx 2^{B_{int}} n_Q P_{inv} f_s \quad (10.38)$$

Circuit nonidealities due to the ADC are greatly suppressed by the modulator loop gain, whereas the digital-analog conversion (DAC) errors are not attenuated. These DAC errors are directly input-referred and appear at the modulator output as additional noise and distortion [39]. Thus, the DAC linearity has to be as high as the overall modulator resolution. To ensure the desired modulator resolution, many DAC linearisation techniques have been

proposed [29] [38]. All of these methods consume additional power and silicon area. For reasons of simplicity, we neglect these factors in the *FOM* calculation for resolution up to 14bits, because it is possible to design DACs with such a linearity [40], [41].

10.5.1.3 Cascade Architectures

Cascaded or multi-stage $\Sigma\Delta$ modulators consist of a cascade connection of low-order modulators $M \leq 2$ (see Fig. 10.4) which are inherently stable [17]. The quantization noise generated in the first stage will be remodulated in the next one, and later cancelled by digital filters. Hence, the characterisation of the digital power consumption in Eq. 10.38 has to be adapted. The number of quantizers is proportional to the number of modulator loops n_{loop} . Furthermore, the additional filters, which remove the quantization errors of the first stages, have to be considered [17]. A coarse approximation of the additional hardware effort for different modulator orders and structures is shown in Tab. 10.4.

Table 10.4 Additional hardware effort in multi-loop $\Sigma\Delta$ modulators.

Structure	Order M	n_{loop}	m_{Adder}	m_{Delay}
2-1 modulator	3	2	8	23
2-2 modulator	4	2	9	23
2-1-1 modulator	4	3	21	62

n_{Adder} and n_{Delay} represent the size of the adders and flip-flops converted in CMOS gate-equivalent inverters, whereas m_{Adder} and m_{Delay} specify the number of adders and flip-flops. Hence, the adjusted digital power consumption results to Eq. 10.39:

$$P_{digital}^{**} = (n_Q * m_{loop} + n_{Adder} * m_{Adder} + n_{Delay} * m_{Delay}) P_{inv} f_s \quad (10.39)$$

Other architectures with only a first-order modulator in the first stage are not interesting in practise, due to the fact that they show a large sensitivity to finite integrator gain and coefficient mismatch.

10.5.2 Design Example: A 10bit – 25kHz CT $\Sigma\Delta$ Modulator

The objective of this section is to find the power efficient modulator implementation for a given set of specifications. We assume in the following that the amplifier is a folded-cascode OTA, with a total current of $4I_B$. The supply voltage is 1.5V with a resulting full scale input amplitude of $\hat{V}_{sw} = V_{DD}/3$. The used 3.3V, $0.5\mu m$ CMOS technology

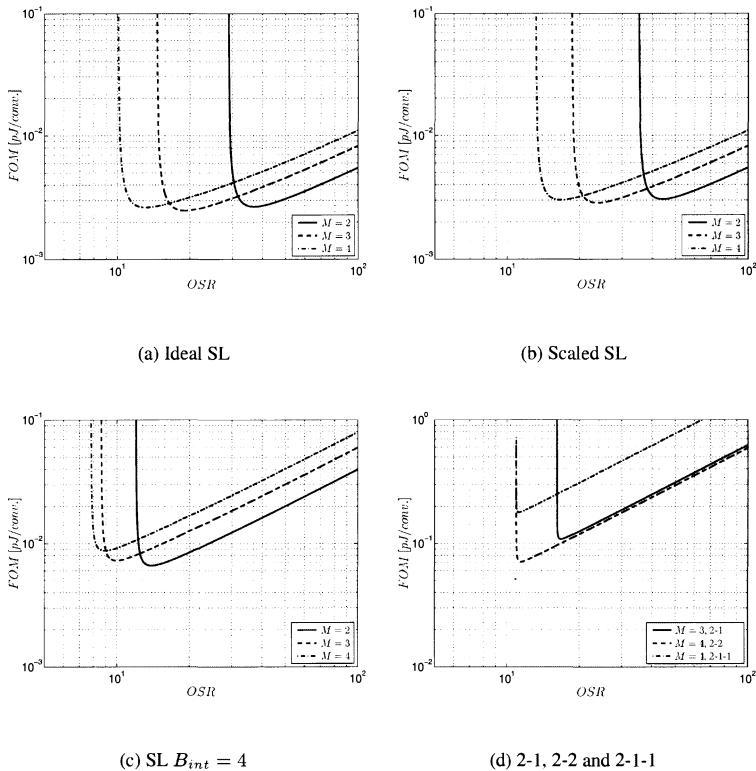


Figure 10.24: Estimated *FOM* as a function of the oversampling ratio *OSR* corresponding to the design example (*10bit – 25kHz*) for different modulator topologies and orders.

provides a transconductance parameter β_{0p} of $40\mu A/V^2$.

The *FOM* estimations will be done to guarantee a modulator resolution of $B = 10bit$ in a signal bandwidth of $25kHz$. Figure 10.24(a) illustrates the *FOM* for a single-loop (SL) single-bit quantizer according to Eq. 10.30 and 10.35. The best *FOM* is obtained with a third-order modulator with an OSR close to 18. A more realistic *FOM* estimation takes also the scaling of the noise transfer function into account Eq. 10.36. This results in Figure 10.24(b). The minimal *FOM*, as well as the minimal *OSR* at which the desired performance is achieved, is increased. In other words, the minimal power consumption increases for a fixed resolution and signal bandwidth. Again, the best *FOM* is obtained with a third order modulator with an OSR of approximately 23. Figure 10.24(c)

and 10.24(d) show the estimated *FOM* for single-loop 3-bit modulators with respect to Eq. 10.37 and 10.38 and for multi-loop (2-1, 2-2, 2-1-1) $\Sigma\Delta$ modulators considering Eq. 10.39. As expected, one can see that the required *OSR* is much lower compared with a single-loop single-bit modulator, *but* on the other hand the power efficiency is about 1 decade worse.

10.5.3 Low-Power Third Order $\Sigma\Delta$ Modulator

In a 3rd order $\Sigma\Delta$ modulator the *SNR* is in general limited by circuit noise (see [23]). The complete fully differential circuit implementation is shown in Fig. 10.25.

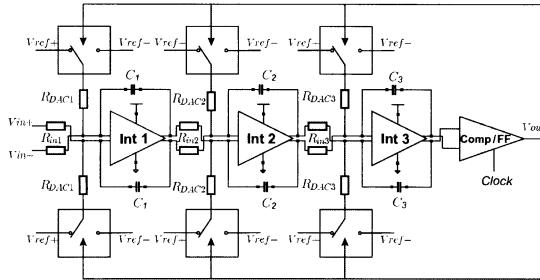


Figure 10.25: Implementation of the CT $\Sigma\Delta$ modulator.

The filter was implemented using three active RC-integrators (**Int 1-3**). In order to combine the high linearity and low-power requirements in the integrators, a single stage folded-cascode amplifier has been chosen [23]. To enhance the output swing of the used amplifier all output transistors have the same low $v_{DS,sat} = 100mV$, whereas the input transistors are biased slightly in weak inversion with $v_{DS,sat} \approx 0$ [23]. Consequently, the input signals of the integrators are referenced to $650mV$. Furthermore, the positive reference voltage ($Vref+$) is set to $975mV$ whereas the negative reference voltage ($Vref-$) is equal to $325mV$.

Depending on the output state of the comparator the positive or negative reference voltage is switched to the feedback resistors (R_{DAC}). The sampling frequency of the A/D converter is set at $2.4MHz$. The used scaling coefficients for the NRZ-DAC are $k_{1_{NRZ}} = 0.05$, $k_{2_{NRZ}} = 0.3$ and $k_{3_{NRZ}} = 0.6416$ for the first, second and third integrator respectively.

10.5.4 Measurement Results

Fig. 10.26(a) shows the resulting signal-to-noise and distortion ratio (SNDR) versus the normalized input signal $Vin/Vref$. The $\Sigma\Delta$ output bit stream has been recorded with a

logic analyser, so that the data can be processed off line with MATLAB. Accordingly, the fast Fourier transformation (FFT) with 65k points and the Blackman-Harris window have been applied. The dynamic range and the peak SNR is 77 dB and 73dB respectively. This corresponds to a resolution of 12bit.

The $\Sigma\Delta$ modulator was implemented in a 3.3V, $0.5\mu m$ triple-metal standard analog CMOS technology ($V_{Tn} = 0.58mV$, $V_{Tp} = 0.62mV$). The measured IBN, power consumption P and the resulting FOM as a function of the supply voltage is shown in Fig. 10.26(b). The modulator works already fine at a supply voltage of 1.3V without any performance degradation. This illustrates the low voltage capability of the realized CT modulator. Furthermore, it is apparent that the minimal FOM is obtained at a supply voltage of 1.4V and rises corresponding to the overall power consumption P . The power efficiency of various A/D converters with different resolutions and sampling rates can be compared using again the FOM . Fig. 10.26(b) demonstrates that this proposed design strategy results in a good power efficiency, compared to [17]. The required area is less than $1mm^2$ with the first integrator consuming more than 50% of the total power consumption and area.

10.6 Conclusion

$\Sigma\Delta$ modulators provide a favorable way of implementing A/D converters for a wide range of sampling frequencies and resolutions. Due to the common demand for higher

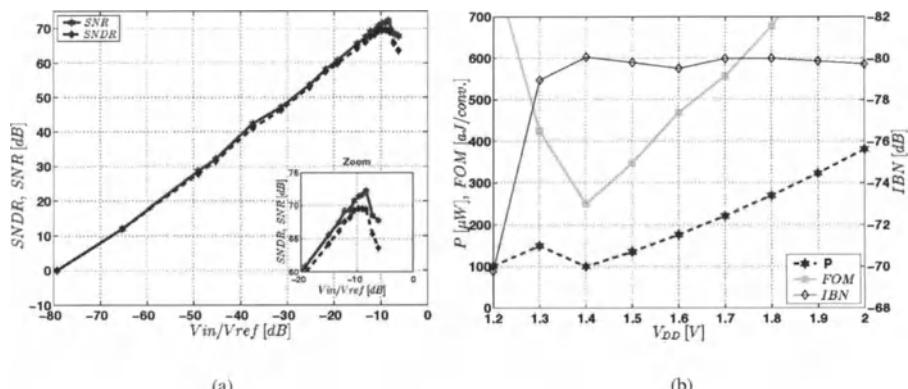


Figure 10.26: (a) Measured signal-to-noise and distortion ratio and (b) measured IBN, power consumption P and the resulting FOM as a function of the supply voltage.

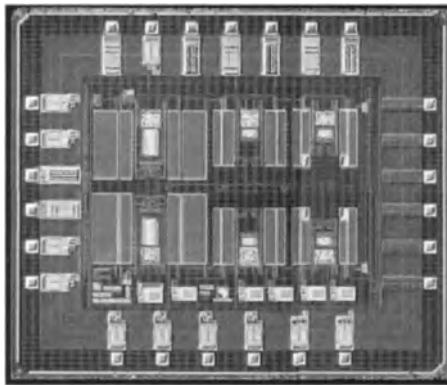


Figure 10.27: Chip photo of the implemented CT $\Sigma\Delta$ modulator.

operating speed, lower power consumption and the use of low supply voltages, the use of continuous-time $\Sigma\Delta$ modulators has become very popular over the last years. In this chapter some of the most recent work on CT $\Sigma\Delta$ modulators has been presented, which can be used in either low pass or band pass applications. First the main focus was set on the equivalence of DT and CT modulator performance as well as the typical continuous-time modulator non-idealities, whose influence have been studied analytically and by simulations.

Next, different possibilities have been presented for preventing the performance degradation in CT modulators coming from two of the most critical non-idealities, i.e. clock jitter and integrator gain mismatch. Finally, a systematic approach was introduced for the design of low power CT $\Sigma\Delta$ modulators based on a figure of merit, which has been proven by a design example.

References

- [1] J. Doernberg; D. A. Hodges, "A 10 – bit 5 Msample/sec CMOS 2-step flash ADC," *Proc. IEEE Custom Integrated Circuits Conf.*, pp. 18.6/1–18.6/4, 1988.
- [2] R. van de Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters*, Kluwer Academic Pub., 1994.
- [3] S. Northworthy; R. Schreier; G. Temes, *Delta-Sigma Data Converters*, IEEE Press, 1997.
- [4] P. Scholtens; M. Vertregt, " 6b 1.6GSample/s flash ADC in $0.18\mu m$ CMOS using averaging termination , " *Proc. IEEE Int. Solid-State Circuits Conf.*, pp. 128,435,

- 2002.
- [5] R. Schreier; B. Zhang, "Delta-sigma modulators employing continuous-time circuitry," *IEEE Trans. Circuits Syst. I*, vol. 43, no. 4, pp. 324–332, April 1996.
 - [6] O. Shoaei, *Continuous-Time Delta-Sigma A/D Converters for High Speed Applications*, Ph.D. thesis, Carleton University, 1995.
 - [7] J.A. Cherry; W. M. Snelgrove, "Excess Loop Delay in Continuous-Time Delta-Sigma Modulators," *IEEE Trans. Circuits Syst. II*, vol. 46, no. 4, April 1999.
 - [8] J. A. Cherry; W. M. Snelgrove, *Continuous-time Delta-Sigma Modulators for high-speed A/D Conversion*, Kluwer Academic Pub., 1999.
 - [9] E.I. Jury, *Theory and Applications of the Z-Transform Method*, John Wiley & Sons, 1964.
 - [10] H. Aboushady; E. de Lira Mendes; M. Dessouky; P. Loumeau, "A Current-Mode Continuous-time $\Sigma\Delta$ Modulator with delayed Return-to-Zero Feedback," *Proc. IEEE Int. Symp. on Circuits and Syst.*, vol. 1, pp. XX, Sep. 1999.
 - [11] F. Gerfers; Y. Manoli., "A 1.5V Low-Power Third Order Continuous-Time Lowpass $\Sigma\Delta$ A/D Converters," *Proc. Int. Symp. on Low-Power Electronics and Design*, pp. 219 –221, July 2000.
 - [12] M. Ortmanns; F. Gerfers; Y. Manoli, "Jitter insensitive Feedback DAC for Continuous-Time $\Sigma\Delta$ Modulators," *Proc. IEEE Int. Conf. on Electronics, Circuits and Syst.*, pp. 1049–1052, 2001.
 - [13] M. Ortmanns; F. Gerfers; Y. Manoli, "On the synthesis of cascaded continuous-time $\Sigma\Delta$ Modulators," *Proc. IEEE Int. Sym. on Circuits and Syst.*, pp. 419–422, 2001.
 - [14] G. Fischer; A. J. Davis, "Alternative topologies for sigma-delta modulators-a comparative study," *IEEE Trans. Circuits Syst. II*, vol. 44, no. 10, pp. 789 –797, Oct. 1997.
 - [15] A. Marques; V. Peluso; M. S. Steyaert; W. M. Sansen, "Optimal Parameters for $\Delta\Sigma$ Modulator Topologies," *IEEE Trans. Circuits Syst. II*, vol. 45, pp. 1232–1241, Sep. 1998.
 - [16] C.H. Lin; M. Ismail, "Synthesis and Analysis of High-Order Cascaded Continuous Time $\Sigma\Delta$ Modulators," in *Proc. Int. Conf. on Electronics, Circuits and Systems*, 1999, pp. 1693–1696.
 - [17] F. Medeiro; B. Perez-Verdu; A. Rodriguez-Vazquez, *Top-Down Design of high-performance Sigma-Delta Modulators*, Kluwer Academic Pub., 1999.
 - [18] F. Gerfers; M. Ortmanns; Y. Manoli, "A Design Strategy for Low-Voltage Low-Power Continuous-Time $\Sigma\Delta$ A/D Converters," *Design, Automation and Test Conf.*, pp. 361–368, 2001.
 - [19] J.F. Jensen; G. Raghavan; A.E. Cosand; R.H. Walden, "A 3.2-ghz second-order delta-

- sigma modulator implemented in inp hbt technology," *IEEE J. Solid-State Circuits*, vol. 30, no. 10, pp. 1119–1127, Oct. 1995.
- [20] M. Ortmanns; F. Gerfers; Y. Manoli, "Successful Design of cascaded Continuous-time $\Sigma\Delta$ Modulators," *Proc. IEEE Int. Conf. on Electronics, Circuits and Syst.*, pp. 321–324, 2001.
 - [21] N. Wongkomet, *A Comparison of Continuous-Time and Discrete-Time Sigma-Delta Modulators, Master Thesis*, University of California at Berkeley, 1995.
 - [22] C. C. Enz; G. C. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: Autozeroing, correlated double sampling and chopper stabilization," *Proc. of the IEEE*, vol. 84, no. 11, pp. 1584–1612, Nov. 1996.
 - [23] Friedel Gerfers; M. Ortmanns; Y. Manoli, "A 12-Bit Power Efficient Continuous-Time $\Sigma\Delta$ Modulator with $220\mu\text{W}$ Power Consumption," *Proc. European Solid-State Circuits Conf.*, pp. 536–539, 2001.
 - [24] B. Razavi, *Design of Analog CMOS Integrated Circuits*, McGraw-Hill, 1999.
 - [25] J.A. Cherry; W.M. Snelgrove, "Clock jitter and quantizer metastability in continuous-time delta-sigma modulators," *IEEE Trans. Circuits Syst. II*, vol. 46, 1999.
 - [26] H. Tao; L. Toth; J.M. Khouri, "Analysis of timing jitter in bandpass sigma-delta modulators," *IEEE Trans. Circuits Syst. II*, vol. 46, 1999.
 - [27] O. Oliae; H. Aboushady, "Jitter Effects in Continuous Time $\Sigma\Delta$ Modulators with delayed Return-To-Zero Feedback," in *Proc. Int. Conf. on Electronics, Circuits and Systems*, 1998, pp. 351–354.
 - [28] B.E.Boser; B.A.Wooley, "The design of sigma-delta modulation analog-to-digital converters," *IEEE J. Solid-State Circuits*, vol. 23, no. 6, pp. 1298–1308, Dec. 1988.
 - [29] R. E. Radke; A. Eshraghi; T. S. Fiez., "A 14-B, Current-Mode $\Sigma\Delta$ DAC Based Upon Rotated Data Weighted Averaging," *IEEE J. Solid-State Circuits*, vol. 35, no. 8, August 2000.
 - [30] Y. Matsuya; N. Yukhi; Y. Akazawa, "Digital Correction Technique for Multi-Stage Noise-Shaping with an RC-Analog Integrator," *IEICE Trans. Elektron*, vol. E77-C, no. 12, pp. 1912–1919, 1994.
 - [31] X. Huang, *Investigation of a Switched-Capacitor concept for Feedback DAC in high resolution CT $\Sigma\Delta$ modulators, Master Thesis*, University of Saarland, Germany, 2000.
 - [32] M. Ortmanns; F. Gerfers; Y. Manoli, "A Continuous-Time Sigma-Delta Modulator with Reduced Jitter Sensitivity," in *Proc. European Solid-State Circuits Conf.*, 2002, p. to be published.
 - [33] F. Gerfers; K. M. Soh; M. Ortmanns; Y. Manoli, "Figure of Merit based Design Strategy for low-power Continuous-Time $\Sigma\Delta$ Modulators," *Proc. IEEE Int. Sym. on*

- Circuits and Syst.*, pp. 233–236, 2002.
- [34] Rudy J. Van De Plassche, “A Sigma-Delta Modulator as an A/D Converter,” *IEEE Trans. Circuits Syst.*, vol. 25, no. 7, pp. 510–514, Juli 1978.
 - [35] J. C. Candy, “A use of double integration in sigma-delta modulation,” *IEEE Trans. on Communications*, vol. COM-33, pp. 249–258, Mar. 1985.
 - [36] R. Schreier; M. Snelgrove, “Stability in a sigma-delta modulator,” *Int. Conference Acoustics, Speech, Signal Processing*, vol. 3, pp. 1769 –1772, 1991.
 - [37] Rex T. Baird; Terri S. Fiez, “A Low Oversampling Ratio 14-b 500-kHz $\Delta\Sigma$ ADC with a Self-Calibrated Multibit DAC,” *IEEE J. Solid-State Circuits*, vol. 31, no. 3, pp. 312–320, März 1996.
 - [38] Y. Geerts; M. S. J. Steyaert; W. M. Sansen, “A high-performance multibit $\Delta\Sigma$ CMOS ADC,” *IEEE J. Solid-State Circuits*, vol. 35, pp. 1829–1840, Dec. 2000.
 - [39] F. Medeiro; B. Perez-Verdu; J. M. de la Rosa; A. Rodriguez-Vazquez, “Fourth-order cascade sc $\sigma\delta$ modulators: A comparative study,” *IEEE Trans. Circuits Syst. I*, vol. 45, no. 10, pp. 1041–1051, Oct. 1998.
 - [40] J. Bastos; A. M. Marques; M.J. Steyaert; W. Sansen., “A 12-Bit Intrinsic Accuracy High-Speed CMOS DAC,” *IEEE J. Solid-State Circuits*, vol. 33, no. 12, Dec. 1998.
 - [41] G. Van der Plas; J. Vandebussche; W. Sansen; M. Steyaert; G. Gielen., “A 14-Bit Intrinsic Accuracy Q^2 Random Walk CMOS DAC,” *IEEE J. Solid-State Circuits*, vol. 34, no. 12, Dec. 1999.

Chapter 11

Bandpass Sigma-Delta A/D Converters: Fundamentals, Architectures and Circuits

*José M. de la Rosa, Belén Pérez-Verdú, Rocío del Río,
Fernando Medeiro and Angel Rodríguez-Vázquez*

11.1 INTRODUCTION^{†1}

During the last years we have witnessed an unprecedented growth of the wireless communication market, which has resulted in the proliferation of many portable appliances such as cellular phones, digital radio receivers, Global Positioning System (GPS) handheld units, Personal Digital Assistants (PDAs), etc. Besides, new technologies and applications are continuously arising such as Bluetooth, Software-Defined Radios (SDRs), Wireless Local Area Networks (WLANS), etc.

The *wireless revolution* has been partially prompted by the ‘vertiginous’ scaling-down of microelectronic technologies, which has exponentially increased the capabilities of digital and mixed-signal VLSI circuits, making it possible the integration of entire communication systems on a single silicon substrate. Together with reduced price, size and power consumption, the so-called Systems-on-Chips (SoCs) tend to realize more and more functions using Digital Signal Processors (DSP), thus facilitating the *programmability* and adaptability of these functions to different radio interface standards [1] – preferably through software reconfiguration [2][3]. For this purpose it would be desirable to place the analog-to-digital interface as closer to the antenna as possible.

Analog-to-Digital Converters (ADCs) based on BandPass $\Sigma\Delta$ Modulators (BP $\Sigma\Delta$ Ms) are very well suited for the implementation of the front-end of such digital wireless communication SoCs. One of the main advantages of BP $\Sigma\Delta$ Ms as compared to other architectures comes from the fact that they do not need to digi-

1 . **Acknowledgement:** This work has been partially funded by the spanish MCyT and the ERDF (Project TIC2001-0929-ADAVERE), the EC IST Project 29261 (MIXMODEST), and the EC-IST Project 2001-34283 (TAMES).

itize the whole Nyquist band (from DC to one half of the sampling frequency); instead they digitize just the signal band, thereby requiring much less power consumption to obtain similar dynamic range.

The principle of $\Sigma\Delta$ Modulation ($\Sigma\Delta M$) [4] is extended in BP $\Sigma\Delta M$ s to bandpass signals, especially but not only, with a narrow bandwidth [5]. Thus, BP $\Sigma\Delta M$ s have much in common with their lowpass counterparts – whose properties have been covered in previous Chapters of this book. However, there are some issues which are peculiar to BP $\Sigma\Delta$ -ADCs. This Chapter is devoted to the description of these issues. In Section 11.2, digital radio receivers are revised, pointing out the necessity for an ADC at the IF location. Section 11.3 and Section 11.4 explain the basic concepts and architectural issues of BP $\Sigma\Delta$ -ADCs. The problems derived from circuit implementation are treated in Section 11.5. Finally, Section 11.6 summarizes the performance of state-of-the-art BP $\Sigma\Delta$ -ADCs.

11.2 THE DIGITAL WIRELESS COMMUNICATIONS UNIVERSE

Since the invention of the cellular concept by Bell Laboratories and the later introduction of the 1st-Generation (1G) (analog) mobile phones in the early 80's, a multitude of wireless communication systems have appeared. As an illustration, Table 12.1 summarizes some significant wireless systems including the 2G and 3G cellular standards as well as others like Bluetooth and WLAN [6]. The table shows the main characteristics of these systems, namely: frequency range, channel spacing, multiple access technique and modulation type. Among other 2G standards, Global System for Mobile communications (GSM) is the most used worldwide. In the case of 3G systems, as they are in their final stage of standard-

Table 11.1: Summary of digital wireless communication standards.

Wireless Standard	Type	Access	Modulation Type	Frequency Range (Tx) (MHz)	Frequency Range (Rx) (MHz)	Channel Spacing (kHz)	Data Rate (kb/s)	Peak Power (W)
AMPS	Cellular1G	FDD	FM	824-849	869-894	30	N/A	3
GSM	Cellular2G	TD/FDMA /FDD	GMSK	890-915	935-960	200	270.8	0.8-8
IS-54		TDMA /FDD	$\pi/4$ -QPSK	824-849	869-894	30	48	0.8-3
IS-95		CDMA	OQPSK	824-849	869-894	1250	1228	N/A
UMTS	Cellular3G	CDMA	QPSK	1920-1980	2110-2170	5000	3840	0.125-2
DCS-1800	Cordless	TDMA	GMSK	1710-1785	1805-1850	200	270.8	0.8-8
DECT		TDMA /TDD	GFSK	1881-1897	1881-1897	1728	1152	0.25
PCS-1900	Wireless Digital	TDMA	GMSK	1880-1910	1930-1955	200	270.8	0.8-8
DSSS		CDMA	QPSK	2400-2483	2400-2483	N/A	1211	1
Bluetooth		CDMA/FH	GFSK	2400-2483	2400-2483	1000	1000	20dBm

ization, there is still an uncertainty about what protocols will dominate the market.

This *wireless communications universe* calls for the need of developing digital transceiver chips capable of operating at different modes with multi-standard support features. The final goal is to allow software-based radios to simultaneously carry voice, video and data, using a variety of telecommunication systems [3].

11.2.1 The ideal digital wireless transceiver

Fig. 11.1 shows the generic block diagram of an ideal digital RF transceiver intended for SDRs. At the receiver side, the RF signal coming from the antenna is directly digitized by an ADC. Hence, many functions like frequency tuning and translation, filtering, channel selection and demodulation can be implemented by running software on a general-purpose DSP. A similar reasoning holds at the transmitter side. Unfortunately, the transceiver of Fig. 11.1 is, to say the least, hard to implement because the specifications required for both the ADC and the DAC are very demanding. For instance, the receiver would require performing analog-to-digital conversion on signals located at 800MHz-2.5GHz (see Table 12.1) with 14-18bit accuracy [7]. Hence, a more realistic digital radio transceiver would contain an Analog Signal Processing (ASP) section including signal conditioning, i.e: frequency translation, amplification and filtering. The implementation of these analog functions can be realized in different manners resulting in several Radio Frequency (RF) transceiver architectures. In this Chapter, we are interested in the receiver architectures – analyzed in the following section.

11.2.2 Overview of wireless digital radio receiver architectures

Fig. 11.2 shows the simplified block diagram of the most significant digital radio receiver architectures [8]. Fig. 11.2(a) is a *digital superheterodyne* receiver, where the signal is first down-translated to an Intermediate Frequency (IF) – for example 10.7MHz in GSM systems – and then to baseband where it is digitized and demodulated. This is the most conventional architecture. However, it is not appropriate for fully-integrated RF receivers because external narrowband filters

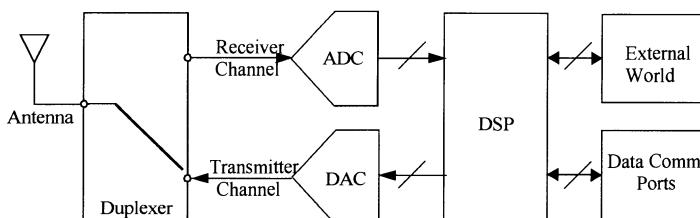


Figure 11.1: Ideal wireless transceiver intended for software-defined radios.

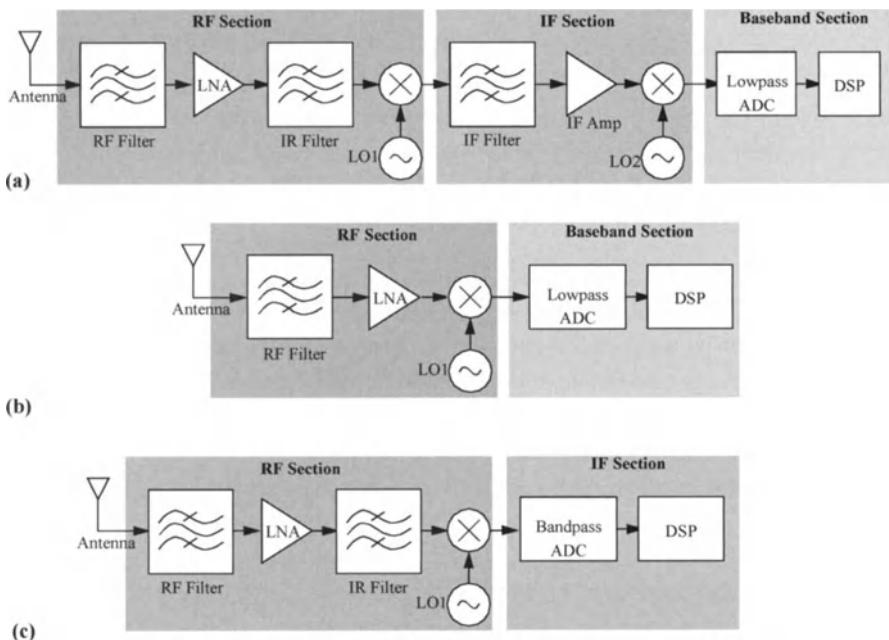


Figure 11.2: Digital RF receivers. (a) Superheterodyne. (b) Direct conversion.
(c) IF conversion.

with high selectivity, such as ceramic or SAW filters (in both the RF and the IF sections) are required. This problem can be partially solved by using a sufficiently low IF – typically $< 1\text{MHz}$ – referred to as *low-IF superheterodyne receivers*. In this case the channel-select filters can be implemented using established circuit techniques such as $g_m - C$ or Switched-Capacitor (SC). However, the use of a low IF imposes very demanding requirements for the Image-Reject (IR) filter².

In order to *relax* the IR filter requirements, an analog quadrature mixer can be used in both the RF and the IF sections [9]. As an illustration, Fig. 11.3(a) shows the IF section of a superheterodyne receiver including a quadrature mixer. Note that two lowpass ADCs are needed to digitize the resulting In-phase (I) and Quadrature (Q) components. These quadrature signals are complex combined in order to cancel the image power in the downconverted signal band. In practice, however, there is not a complete cancellation of the image due to gain and phase mismatch between quadrature paths in Fig. 11.3(a).

For that reason, other alternatives have been explored in last years to improve the IR problem in *low-IF superheterodyne receivers*. One of them is based on the

2 . The problem of suppressing signals in the image band will be discussed later.

integration of an IF quadrature mixer with a lowpass $\Sigma\Delta M$, usually referred to as IF-to-Baseband $\Sigma\Delta M$ s [10][11]. In these architectures, the mixer errors are shaped so that they are reduced in the desired band. In practical circuit realizations, the IR feature is limited by I/Q path mismatches, which requires the use of additional compensation strategies like dynamic element matching algorithms [10].

The IR problem can be completely eliminated by using the receiver shown in Fig. 11.2(b), known as *direct conversion, zero-IF* or *homodyne* receiver. In this architecture, the RF signal is mixed-down directly to DC where it is digitized. This approach is more suited to integration than the *superheterodyne* because it eliminates the IF section. Hence, only off-chip RF filters are required since, as the image and the desired signal are the same, the IR filter can be removed. However, the offset and flicker noise of the mixer are present in the middle of the signal band and can severely degrade the performance of this type of receivers.

Many of the problems arising in the above mentioned receiver architectures can be eliminated using the *IF-conversion receiver*, shown in Fig. 11.2(c). In this architecture the in-coming signal at the antenna is first mixed-down to IF where it is digitized. Thus, the signal is first translated to the digital domain by one ADC, referred to as *bandpass* or *IF ADC*, and then mixed to the baseband as shown in Fig. 11.3(b). This is advantageous for several reasons. On the one hand, as quadrature mixing is done in the digital domain, the problems associated with the analog mixer in the receiver shown in Fig. 11.3(a) are avoided [12]. Another advantage of the *IF-conversion receiver* is that it allows channel-select filtering, gain control and demodulation to be handled in the digital domain [13][14]. This results in robust RF receivers with a high degree of programmability, thus allowing a single software-controlled RF receiver to be employed for multi-standard receivers [15].

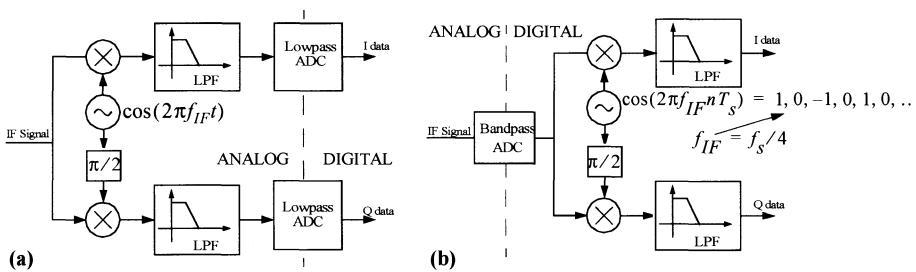


Figure 11.3: Using quadrature mixers in: (a) Superheterodyne receivers. (b) IF-conversion receivers.

11.2.3 IF A/D Conversion – Bandpass $\Sigma\Delta$ modulators

Digitization in IF-conversion receivers can be accomplished either with a wideband Nyquist-rate ADC or a BP $\Sigma\Delta$ -ADC. The use of the latter is the optimum solution since the bandwidth of IF signals is typically much smaller than the carrier frequency, and hence, reducing the quantization noise in the entire Nyquist band becomes superfluous. Instead, by using BP $\Sigma\Delta$ -ADCs the quantization noise power is reduced only in a narrowband around the IF location, thus taking advantage of the higher *oversampling ratio*^{†3} and hence yielding to a high resolution.

BandPass $\Sigma\Delta$ Modulators (BP $\Sigma\Delta$ Ms) extend the noise-shaping concept from the *conventional LowPass $\Sigma\Delta$ Ms* (LP $\Sigma\Delta$ Ms) – in which the quantization noise is suppressed around DC – to a more general case where the quantization noise is reduced in a narrow passband centred at an IF location [5][17]. Thus, the design and analysis of BP $\Sigma\Delta$ Ms share much in common with LP $\Sigma\Delta$ Ms [12].

The rest of the chapter is devoted to the study of BP $\Sigma\Delta$ -ADCs, surveying the different design issues, from architectures to circuit implementation.

11.3 BASIC CONCEPTS OF BANDPASS $\Sigma\Delta$ A/D CONVERTERS

A BP $\Sigma\Delta$ ADC^{†4} is a particular class of $\Sigma\Delta$ -ADC that places the zeroes of the quantization noise transfer function in a narrow band around an IF location, usually named *notch* frequency, and represented by the parameter f_n .

Fig. 11.4 shows the conceptual block diagram of a BP $\Sigma\Delta$ M. It is composed of a BandPass Filter (BPF), an N -bit quantizer and a Digital-to-Analog Converter (DAC) connected in a loop. The BPF can be synthesized by cascading two or more second-order biquadratic filters or *resonators*, which must have a sharp transfer function and well-defined resonance at f_n . These resonators may be implemented as a Discrete-Time (DT) filter using either SC [13][14] or Switched-current (SI) [18] techniques or they may be implemented as a Continuous-Time (CT) filter [19][20].

3 . In the case of bandpass signals, the oversampling ratio is defined as [16]

$M = f_S \lfloor (f_n + B_w/2)/B_w \rfloor / [2(f_n + B_w/2)]$, where B_w is the signal bandwidth, f_S is the sampling frequency, f_n is the centre frequency and $\lfloor x \rfloor$ stands for the largest integer not exceeding x .

4 . A BP $\Sigma\Delta$ ADC is composed of three basic blocks: an anti-aliasing filtering, a BP $\Sigma\Delta$ M and a digital decimator. As in the lowpass case, the modulator is the hardest to design since oversampling simplifies the anti-aliasing filter requirements and the decimator is a pure digital block whose design can be highly structured and automated [12]. For that reason, we will focus on the modulator design issues.

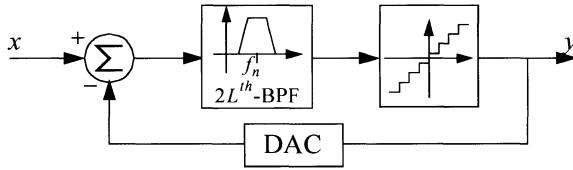


Figure 11.4: Conceptual block diagram of a BPΣΔM.

Let us consider that the BPF is a $2L^{th}$ -order filter composed of a cascade of L resonators with a DT^{†5} transfer function given by:

$$H_R(z) = \frac{N_R(z)}{(1 - z^{-1}z_n)(1 - z^{-1}z_n^*)} \quad (11.1)$$

where z_n and z_n^* are the conjugate-complex poles of $H_R(z)$. Assuming that the quantization error can be modelled as an additive, white noise source, the z -transform of the modulator output in Fig. 11.4 can be written as:

$$Y(z) = S_{TF}(z)X(z) + N_{TF}(z)E_q(z) \quad (11.2)$$

where the signal transfer function and the noise transfer function are respectively:

$$S_{TF}(z) = \frac{[N_R(z)]^L}{[N_R(z) + (1 - z^{-1}z_n)(1 - z^{-1}z_n^*)]^L} \quad (11.3)$$

$$N_{TF}(z) = \frac{[(1 - z^{-1}z_n)(1 - z^{-1}z_n^*)]^L}{[N_R(z) + (1 - z^{-1}z_n)(1 - z^{-1}z_n^*)]^L} \quad (11.4)$$

Note that $N_{TF}(z)$ has L zeroes at $z = z_n$ and $z = z_n^*$. In most practical cases, z_n and z_n^* are placed in the unit circle, i.e., $z_n = \exp[j(2\pi f_n T_S)]$, with T_S being the sampling period. In some BPΣΔMs the value of f_n can be either digitally [21] or continuously [22] programmable, thus allowing f_n to be changed without changing f_S . This is especially useful in radio applications, where the use of a tunable BPΣΔM eliminates the necessity of a channel-selection function in RF receivers.

Assuming that $N_R(z)$ is synthesized in such a way that

5 . A similar discussion can be held for CT-BPΣΔMs. This class of modulators are internally DT systems as will be discussed in Section 11.4.5.

$$N_R(z) + (1 - z^{-1}z_n)(1 - z^{-1}z_n^*) = 1 \quad (11.5)$$

yields:

$$N_{TF}(z) = [1 - 2\cos(2\pi f_n T_S)z^{-1} + z^{-2}]^L \quad (11.6)$$

and the Power Spectral Density (PSD) of the *shaped* quantization noise is:

$$S_Q(f) = \frac{\Delta^2}{12f_S} |N_{TF}(f)|^2 = \frac{\Delta^2}{12f_S} |4\sin[\pi(f-f_n)T_S]\sin[\pi(f+f_n)T_S]|^{2L} \quad (11.7)$$

where Δ is the quantization step, defined as $\Delta \equiv X_{FS}/(2^N - 1)$, with X_{FS} being the full-scale range of the quantizer.

The quantization noise in-band power can be calculated as follows:

$$P_Q = \int_{f_n - B_w/2}^{f_n + B_w/2} 2S_Q(f)df \equiv \frac{(\sin[2\pi f_n T_S])^{2L} \pi^{2L} X_{FS}^2}{12(2^N - 1)^2 (2L + 1) M^{(2L + 1)}} \quad (11.8)$$

where B_w is the signal bandwidth and $B_w \ll f_n$ has been assumed. An important conclusion is that, although the BPΣΔM of Fig. 11.4 is a $2L$ th-order modulator, the quantization noise is suppressed with L th-order bandstop filtering. In other words, the shaping performed on the quantization noise by a $2L$ th-order BPΣΔM is equivalent to that performed by an L th-order LPΣΔM. As an illustration, Fig. 11.5(a) plots several simulated output spectra of the modulator in Fig. 11.4 for $N = 1$, $f_n = f_S/4$ and different values of L . The input is a sinusoidal signal of a frequency close to f_n and an amplitude $A = A_{REF}/2$, with A_{REF} being the output level of the DAC. As Fig. 11.5(a) shows, the output spectrum of the modulator can be seen as the sum of two components: the input signal spectrum (a vertical line) and the quantization noise spectrum. The form of the shaped

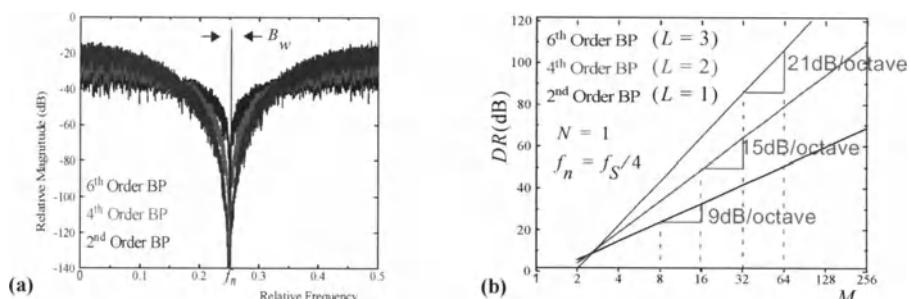


Figure 11.5: (a) Ideal output spectra of a $2L$ th-order BPΣΔM. (b) DR vs. M .

quantization noise is like a valley with its minimum value at f_n .

The Signal-to-Noise Ratio (SNR) and the Dynamic Range (DR) for the modulator of Fig. 11.4, are given respectively by:

$$SNR \equiv \frac{A^2/2}{P_Q} = \frac{3(2^N - 1)^2 (2L + 1) M^{2L+1}}{2\pi^{2L} (\sin[2\pi f_n T_s])^{2L}} \left(\frac{2A}{X_{FS}}\right)^2 \quad (11.9)$$

$$DR \equiv \frac{(X_{FS}/2)^2}{2P_Q} = \frac{3(2^N - 1)^2 (2L + 1) M^{2L+1}}{2\pi^{2L} (\sin[2\pi f_n T_s])^{2L}} \quad (11.10)$$

Note that for a $2L$ th-order BPΣΔM, the DR increases at a rate of $[(2L + 1)\log(2)]$ dB/octave – equivalent to an L th-order LPΣΔM. This is shown in Fig. 11.5(b) by plotting DR vs. M , computed from the spectra in Fig. 11.5(a).

11.3.1 Signal passband location

In theory, the passband of a BPΣΔM can be placed at any frequency from DC to $f_S/2$. Thus, for a given input signal centre frequency, f_{IF} (the IF location in a wireless receiver) and bandwidth, B_w , the choice of the ratio f_n/f_S is a trade-off among sampling frequency (directly related to the speed of the whole system), anti-aliasing filter requirements, and oversampling ratio. As illustrated in Fig. 11.6(a), the transition band, B_{tr} , of the anti-aliasing filter becomes sharper as

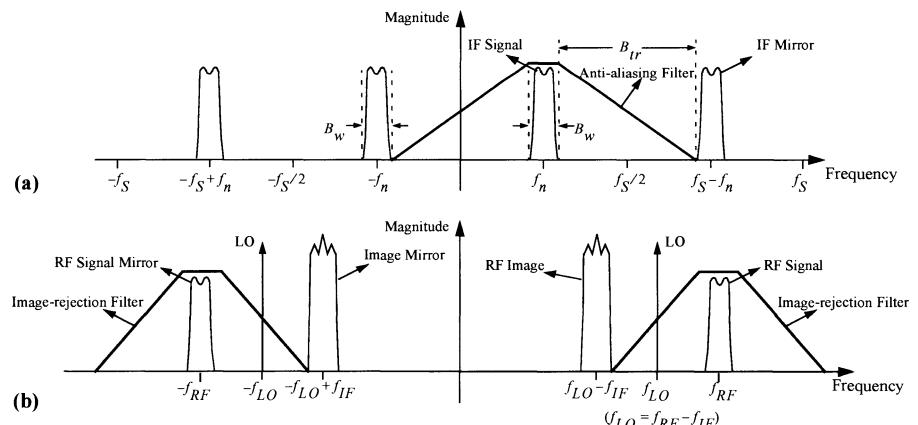


Figure 11.6: Choice of the signal passband location. (a) Trade-off among anti-aliasing filter requirements and high values of f_{IF} . (b) Trade-off among IR filter requirements and low values of f_{IF} .

f_n approaches $f_S'/2$. Thus, the lower $f_n = f_{IF}$ the higher B_{trmax} . However, the use of low f_{IF} complicates the problem of suppressing image-band signals when the RF signal is mixed down to an IF location. This is illustrated in Fig. 11.6(b), where the incoming RF signal is centered at f_{RF} . Note that, an IR filtering must be performed preceding the mixer to avoid image-band signals centered at $f_{RF} - 2f_{IF}$ to corrupt the desired signal [1][8].

In order to cope with both IR and anti-aliasing filter requirements, f_n must be located at an intermediate location in the Nyquist band. An optimum solution to this problem is to place f_n at one-quarter of the sampling frequency. This notch frequency location, in addition to relaxing the mentioned filter specifications, offers several advantages. First, the forward path loop (analog) filter realization can be relatively simplified. Secondly, it makes simpler the synthesis of bandpass architectures, which can be easily derived from lowpass prototypes with a simple variable transformation; for instance

$$z^{-1} \rightarrow -z^{-2} \quad (11.11)$$

as will be described in the next section. Last but not least, the design of the digital mixing to baseband (see Fig. 11.3(b)) is obviously simplified because the digital cosine and sine signals are equal to the data series $(1, 0, -1, 0, \dots)$ and $(0, 1, 0, -1, \dots)$, respectively.

In addition to the mentioned advantages, making $f_n = f_S'/4$ also offers the possibility of centering the IF signal at $3f_S'/4$ as demonstrated in [23] – the spectrum is symmetrical with respect to $f_S'/2$. This approach offers several advantages. On the one hand, making $f_{IF} = 3f_S'/4$ the anti-aliasing filter requirements are the same as for $f_{IF} = f_S'/4$, but the IR filter specifications are relaxed. On the other hand, it allows for either the clock rate to be reduced to one-third or processing signals to be three times higher in frequency. The only drawback is that the oversampling ratio is also reduced by a factor of three. For example, in the case of a 4th-order BPΣΔM, this means a *DR* loss of 3.7bit.

11.3.2 Decimation for bandpass ΣΔ ADCs

The decimator filter is the last stage of a ΣΔ ADC. This block realizes two operations on the modulator output bit stream: filtering the out-of-band quantization noise, and reducing the sampling rate to the Nyquist rate [12].

Fig. 11.7 illustrates the decimation process in BPΣΔ-ADCs. The modulator output, y , is first filtered by a bandpass filter with a digital cut-off frequency at $B_w = f_S'/(2M)$. This filter removes all out-of-band components in order to avoid aliasing in the subsequent *compressor* stage. Thus, the band-limited signal result-

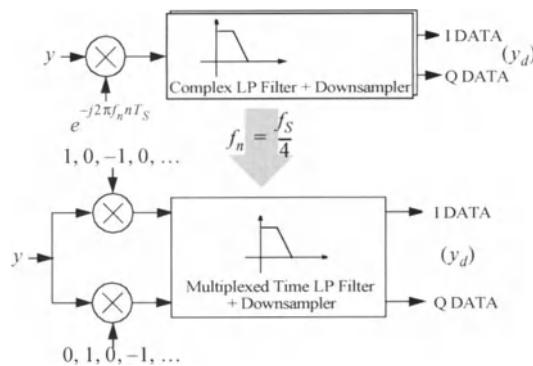


Figure 11.8: Efficient decimator for BP $\Sigma\Delta$ Ms.

ing from the filtering, w , is *downsampled* by discarding $M - 1$ out of every M samples to produce the decimated signal, y_d , at the Nyquist rate.

Note that the scheme of Fig. 11.7 requires a high-Q narrow-band BPF with a high passband center frequency. This yields an increase of cost, in terms of power consumption and silicon area, as compared to the lowpass case. This problem can be avoided by using the scheme shown in Fig. 11.8, composed of a complex mixer and a complex lowpass filter [24]. The modulator output is mixed down to baseband through the multiplication of $\exp(-j2\pi f_n n T_S)$. This scheme can be notably simplified if $f_n = f_S/4$ because the multiplying signal is a sum of two periodic data series containing 0's and ± 1 's as illustrated in Fig. 11.8.

Note that, as the inputs to the lowpass filters are zeros in alternate clock cycles, the two lowpass filters can be simplified by only one multiplexed in time.

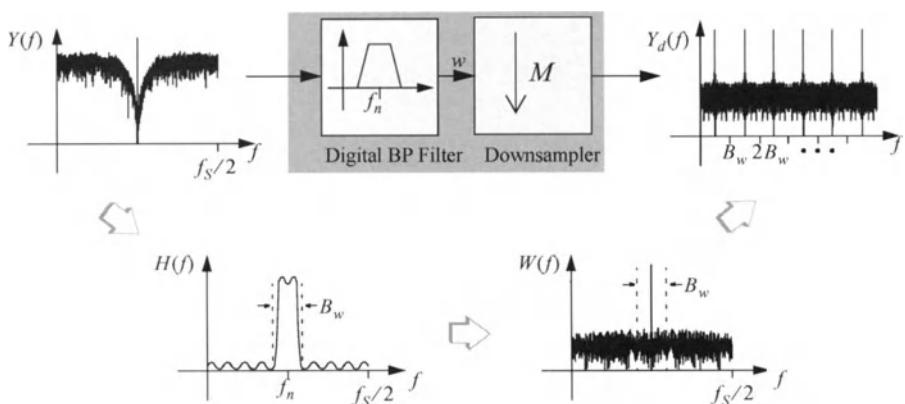


Figure 11.7: Decimation process in BP $\Sigma\Delta$ -ADCs.

11.4 SYNTHESIS OF BANDPASS $\Sigma\Delta$ MODULATOR ARCHITECTURES

The basic structure of a BP $\Sigma\Delta$ M is analogous to that of an LP $\Sigma\Delta$ M except for the type of loop filter. Thus, the operation of both types of $\Sigma\Delta$ Ms is based on the same strategy to attenuate the quantization noise. Hence, although most of the design art developed for LP $\Sigma\Delta$ Ms can be used to develop BP $\Sigma\Delta$ Ms, there are some aspects of the modulator design which are peculiar to BP $\Sigma\Delta$ Ms. This fact has motivated the development of several methods for synthesizing BP $\Sigma\Delta$ M architectures. This section summarizes the most important of these methods.

11.4.1 The lowpass-to-bandpass transformation method: LP-to-BP

As shown in Section 11.3, an L th-order LP $\Sigma\Delta$ M and a $2L$ th-order BP $\Sigma\Delta$ M are equivalent in terms of SNR and DR . Consequently, a simple way to synthesize any BP $\Sigma\Delta$ M architecture is to apply a Lowpass-to-Bandpass (LP-to-BP) transformation to an LP $\Sigma\Delta$ M that meets a given specification. The LP-to-BP transformation most extensively used in BP $\Sigma\Delta$ M Integrated Circuits (ICs) is:

$$z^{-1} \rightarrow -z^{-2} \quad (11.12)$$

Applying the above transformation to the L th-order LP $\Sigma\Delta$ M of Fig. 11.9(b), the $2L$ th-order BP $\Sigma\Delta$ M of the Fig. 11.9(a) is obtained. Assuming a linear model for the quantizer, the Z-transform of the BP $\Sigma\Delta$ M output is given by:

$$Y(z) = (-z^{-2})^L X(z) + (1 + z^{-2})^L E_q(z) \quad (11.13)$$

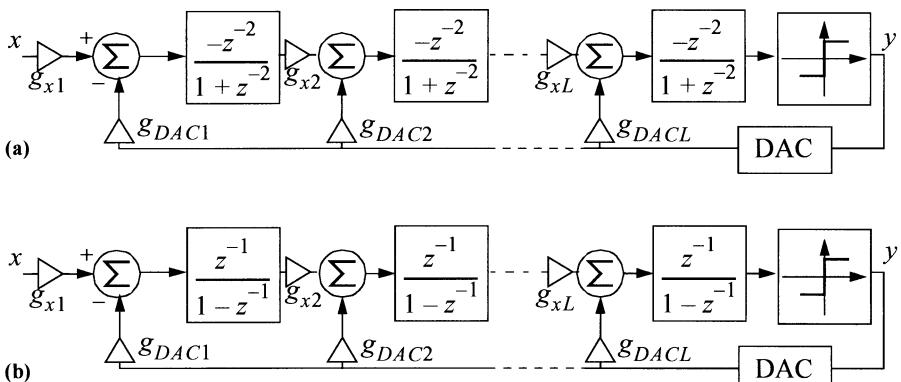


Figure 11.9: (a) Block diagram of the $2L$ th-order BP $\Sigma\Delta$ M derived by applying $z^{-1} \rightarrow -z^{-2}$ to the L th-order LP $\Sigma\Delta$ M shown in (b).

As illustrated in Fig. 11.10, the zeroes of the N_{TF} are mapped from DC (in the original LP $\Sigma\Delta$ M) to $f_S/4$ (in the resulting BP $\Sigma\Delta$ M), which corresponds to $z_n = \exp[j(\pi/2)]$ in Fig. 11.10.

Note that, as a consequence of the transformation in eq.(11.12), the integrators of the original LP $\Sigma\Delta$ M become resonators in the resulting BP $\Sigma\Delta$ M, which have the transfer function in eq.(11.1) with $z_n = \exp[j(\pi/2)]$ and $N_R(z) = -z^{-2}$.

The shaped quantization noise power, P_Q , of the modulator in Fig. 11.9, can be obtained by substituting $f_n = f_S/4$ in eq.(11.8), yielding to an identical expression to that of an L th-order LP $\Sigma\Delta$ M. In an analogous way, the SNR and the DR are identical to that obtained in a L th-order LP $\Sigma\Delta$ M [25].

In general, any $2L$ th-order, N -bit BP $\Sigma\Delta$ M can be obtained by applying the transformation in eq.(11.12) to an L th-order, N -bit LP $\Sigma\Delta$ M. As an illustration, Fig. 11.11 shows several BP $\Sigma\Delta$ Ms obtained by using that transformation. Fig. 11.11(a) is a 2nd-order BP $\Sigma\Delta$ M, Fig. 11.11(b) is a 4th-order, and Fig. 11.11(c) is a 8th-order 4-4 cascade obtained from a 1st-order LP $\Sigma\Delta$ M, a 2nd-order LP $\Sigma\Delta$ M and a 4th-order 2-2 cascade LP $\Sigma\Delta$ M, respectively.

The transformation in eq.(11.12) preserves all features of the original modulator: P_Q , SNR , DR , etc.... In addition to these characteristics, eq.(11.12) keeps the stability properties of the original LP $\Sigma\Delta$ Ms. In fact, the resulting BP $\Sigma\Delta$ M will be stable if and only if the original LP $\Sigma\Delta$ M is stable. As for the lowpass case, the use of cascade BP $\Sigma\Delta$ Ms guarantees the stability for high-order modulators ($L > 2$).

11.4.1.1 Pattern noise of second-order BP $\Sigma\Delta$ Ms

An important property that is inherited from LP $\Sigma\Delta$ Ms is the non-linear behaviour of the quantization error. This phenomenon is much more significant as the number of internal levels of the quantizer and/or the order of the $\Sigma\Delta$ decreases, the worst case corresponding to a 1-bit 1st-order LP $\Sigma\Delta$ M [26]. In the bandpass case, a 1-bit 2nd-order BP $\Sigma\Delta$ M presents a *noise pattern* similar to that of its

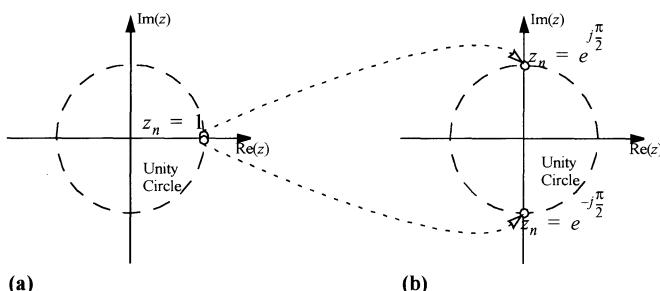


Figure 11.10: Zero location of N_{TF} for: (a) LP $\Sigma\Delta$ M. (b) BP $\Sigma\Delta$ M.

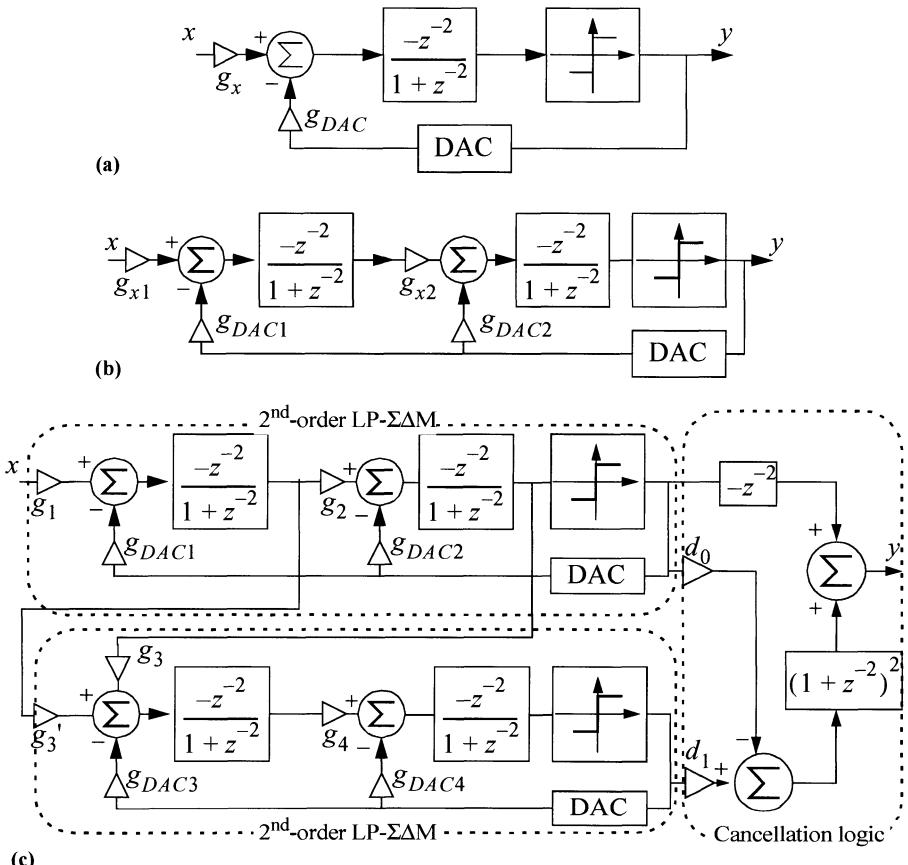


Figure 11.11: BPΣΔMs obtained by using $z^{-1} \rightarrow -z^{-2}$. a) 2th-order BPΣΔM.
b) 4th-order BPΣΔM. c) 8th-order 4-4.

lowpass counterpart. To demonstrate this, the modulator of Fig. 11.11(a) is simulated for a single-tone input signal with a frequency equal to $f_S'/4$. The signal amplitude was varied in the range $[-\Delta/2, \Delta/2]$. Fig. 11.12 (a) shows the quantization error in-band power as a function of the input signal amplitude for $M = 64$, showing a behaviour similar to that shown in 1st-order LPΣΔMs with DC signals [26], reproduced in Fig. 12.12(b).

11.4.1.2 Other LP-to-BP transformations

The transformation in eq.(11.12) places the zeroes of $N_{TF}(z)$ at $f_n = f_S'/4$. This relaxes the IR and anti-aliasing filtering requirements as shown in Section 11.3.1. However, centering the signal passband at $f_S'/4$ has some disadvantages.

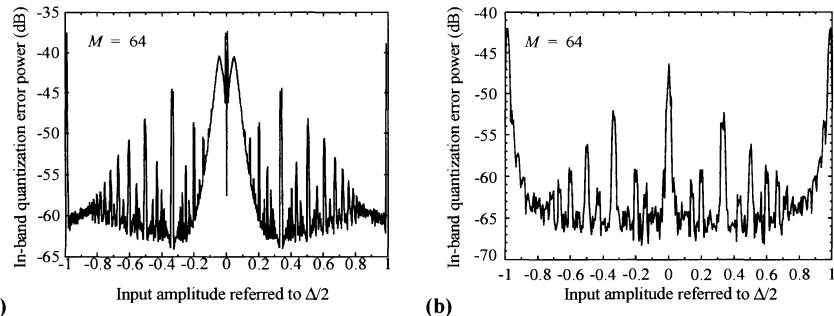


Figure 11.12: Pattern noise of a: (a) 2nd-order BP $\Sigma\Delta$ M. (b) 1st-order LP $\Sigma\Delta$ M.

On the one hand, in the presence of non-linear errors in the analog circuitry of the modulator, any inter-modulation distortion products resulting from the mixing of tones located at $f_S'/2$ with input signal will fall inside the passband, thus corrupting the signal information. This will be treated in more detail in the following sections. On the other hand, for a given input IF, the demands for the sampling rate of the modulator are more restrictive than placing the signal passband center frequency between $f_S'/4$ and $f_S'/2$ [23].

Because the above-mentioned reason, some authors propose to generalize the LP-to-BP transformation in order to locate the zeroes of N_{TF} at any arbitrary frequency, $f_n = f_S'/(2p)$ [27]. However, this approach, generally yielding to complex unpractical architectures, has been rarely used in the reported BP $\Sigma\Delta$ M ICs.

11.4.2 Optimized synthesis of $N_{TF}(z)$

A more flexible approach for designing BP $\Sigma\Delta$ Ms consists on directly synthesizing the modulator loop filter. This allows us to place the poles and zeroes of both $S_{TF}(z)$ and $N_{TF}(z)$ optimally in order to fulfil given specifications [13]. From this perspective, the design of BP $\Sigma\Delta$ Ms is essentially reduced to a problem of filter optimization.

The resulting architectures are usually of the *interpolative* type like that in the lowpass case [12]. This type of architecture offers the possibility of designing $S_{TF}(z)$ in such a way that it performs an anti-aliasing filter. However, as occurs with other interpolative structures, complicated analog circuitry is required, thus being more sensitive to the precision of the components.

For illustration purposes, Fig. 11.13(a) shows a 4th-order BP $\Sigma\Delta$ M, composed of a cascade of resonators, designed using this method [13]. Fig. 11.13(b) shows the pole/zero location of $S_{TF}(z)$ and $N_{TF}(z)$. Observe that the zeroes of $N_{TF}(z)$ are placed around $f_S'/4$. In this case $S_{TF}(z)$ realizes a bandpass function which

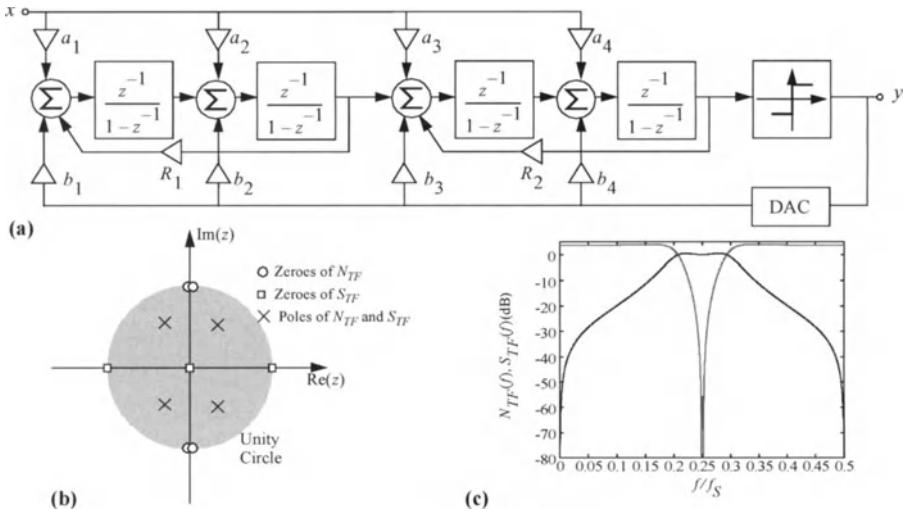


Figure 11.13: An example of a 4th – order BPΣΔM obtained using the directly $N_{TF}(z)$ method [13]. (a) Modulator architecture. (b) Pole/zero location of $N_{TF}(z)$ and $S_{TF}(z)$. (c) $S_{TF}(f)$ and $N_{TF}(f)$ vs. frequency.

acts as an anti-aliasing filter (see Fig. 11.13(c)).

11.4.3 Quadrature bandpass ΣΔ modulators

As stated in Section 11.2, the IR problem arising in digital superheterodyne receivers can be overcome by using quadrature mixers in both the RF and the IF sections. An obvious consequence of using quadrature mixing in the RF section is that the IF signal is separated into two components: I and Q. Hence, two BPΣΔMs are required as illustrated in Fig. 11.14(a), which means doubling the required hardware – two BPΣΔMs compared to only one BPΣΔM if a simple mixer is used. This fact motivates finding new strategies that solve the problem of digitizing the I/Q components of the IF signal.

Fig. 11.14(b) shows an radio architecture that uses a complex, or quadrature, version of a BPΣΔM, called *quadrature* BPΣΔM [28]. This type of BPΣΔMs uses only one ADC to perform directly the A/D conversion of both I and Q signals.

Fig. 11.15 shows a conceptual block diagram of a quadrature BPΣΔM. The main difference with respect to conventional BPΣΔMs is the complex BPF embedded in the loop. Thus, the modulator output consists of a pair of bit streams, one of them representing the real output and the other one the imaginary output.

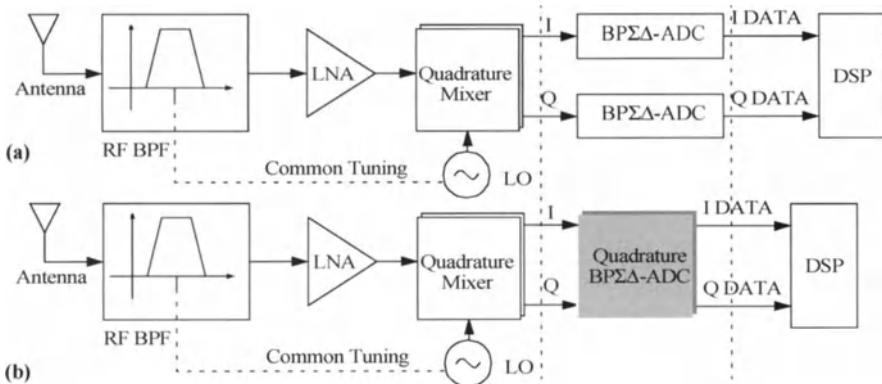


Figure 11.14: RF radio receivers using: (a) Conventional BPΣΔMs. (b) Quadrature BPΣΔMs.

When combined, these two outputs form a complex digital signal which represents the complex (I/Q) input signal and the shaped quantization noise.

The complex BPF in a quadrature BPΣΔM can be realized either using DT or CT circuitry. In practice, this filter is constructed from several cross-coupled real filters as illustrated in Fig. 11.16(a). The complex output signal is:

$$X_o(z) = H(z)X_i(z) = X_{o_{Re}}(z) + jX_{o_{Im}}(z) \quad (11.14)$$

where,

$$\begin{aligned} X_{o_{Re}} &= H_{Re}(z)X_{i_{Re}} - H_{Im}(z)X_{i_{Im}}(z) \\ X_{o_{Im}} &= H_{Re}(z)X_{i_{Im}} + H_{Im}(z)X_{i_{Re}}(z) \end{aligned} \quad (11.15)$$

Observe that there is an analogy between complex filters and fully differential filters in the sense of that both architectures double the number of elements required to implement a given circuit. As an illustration, Fig. 11.16(b) shows a realization of a complex 1st-order filter with a single pole at $z_p = a + jb$ [28].

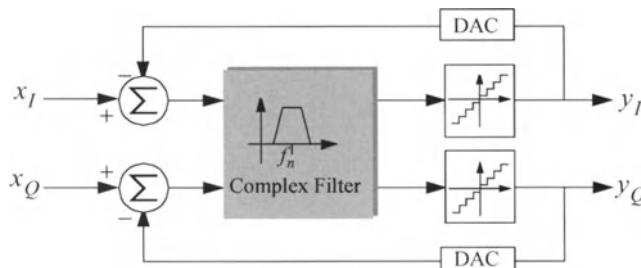


Figure 11.15: Conceptual block diagram of a quadrature BPΣΔM.

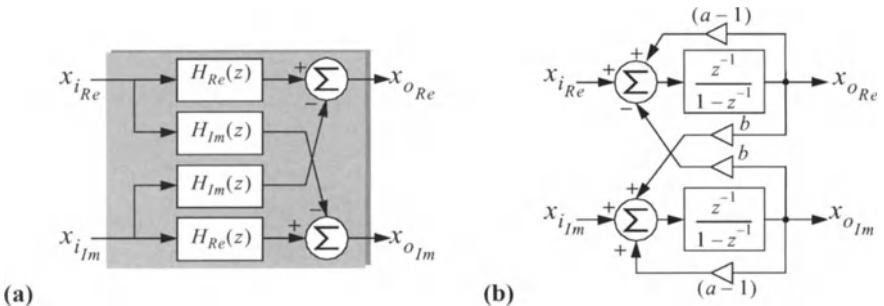


Figure 11.16: Realization of complex filters. (a) Conceptual block diagram. (b) Complex filter with a single pole.

The N_{TF} of quadrature BPΣΔMs has complex value coefficients and hence, it is not constrained to have only complex-conjugate zeroes or to display a symmetric response respect to DC. This allows an L th-order BPΣΔM to place L zeroes at f_n without having any zero at $-f_n$. Therefore, the zeroes of N_{TF} may be a rotated version of those of an L th-order LPΣΔM. To illustrate this, let us consider a complex 4th-order BPΣΔM with,

$$N_{TF} = (1 - jz^{-1})^4 \quad (11.16)$$

This function, displayed in Fig. 11.17, presents four zeroes at $f_n = f_S/4$.

An important practical limitation of quadrature BPΣΔMs is due to mismatching between real and imaginary channels. As a consequence, signal image components will appear in the signal band, thus corrupting the information. To reduce this effect, quadrature BPΣΔMs must be designed to place some of the N_{TF} zeroes at the image band [28]. However, this reduces the order of the quantization noise filtering performed by the quadrature modulator – one of its main advan-

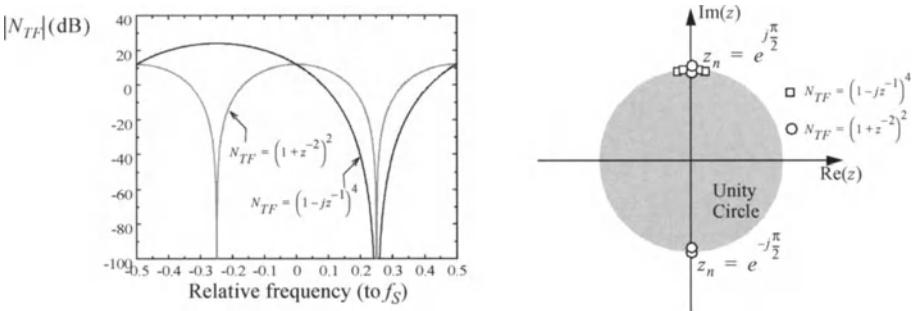


Figure 11.17: N_{TF} and noise zero location for a quadrature 4th-order BPΣΔM.

tages with respect to conventional BPΣΔMs.

11.4.4 N -path bandpass ΣΔ modulators

Centering the notch frequency at $f_S/4$ has multiple advantages as already mentioned. However, in practical applications, the *DR* of BPΣΔMs becomes increasingly constrained by circuit non-idealities at high sampling rates – needed to digitize signals at IF locations. To overcome this problem, some authors propose the use of N -path filters to implement the resonator transfer function [15].

Using the *N-path* design technique, [29], the resonator transfer function can be separated into two high-pass filter, sampled at $f_S/2$, such that:

$$H(z) = \frac{z^{-2}}{1 + z^{-2}} = \frac{z_p^{-1}}{1 + z_p^{-1}} \Bigg|_{z_p = z^2} \quad (11.17)$$

As an illustration, Fig. 11.18 shows a 2-path 4th-order BPΣΔM [30]. The original architecture is partitioned into two interleaved paths, with the resonators replaced by two high-pass filters as in eq.(11.17).

The main problem of N -path architectures is due to the gain and phase mismatches between the different paths. This manifests itself as mirror image signals which appear in the signal bandwidth and corrupt the information [30].

11.4.5 Synthesis of continuous-time bandpass ΣΔ modulators

The architectures described in earlier sections assumed that the loop filter is of the DT type. In recent years, the increased demand for high-speed BPΣΔMs has motivated the development of BPΣΔMs based on CT loop filters, generically known as Continuous-Time BPΣΔMs (CT-BPΣΔMs) [19][20]. This approach offers several advantages. On the one hand, CT filters are much faster than their

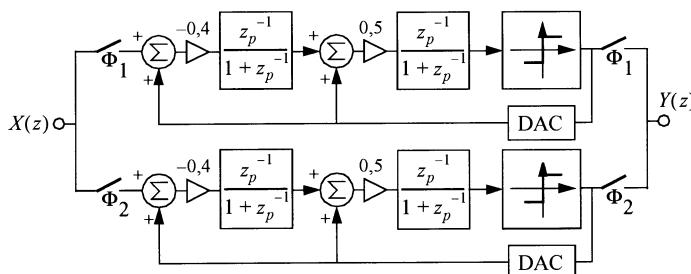


Figure 11.18: Conceptual diagram of a 2-path 4th-order BPΣΔM [30].

DT counterparts. On the other hand, it can be shown that CT-BPΣΔMs provide an implicit anti-aliasing filter for out-of-band signals at no cost [22]. However, CT-BPΣΔMs are more sensitive to *clock jitter* than DT-BPΣΔMs. This is because the internal clock that controls the comparison instant, also controls the rising and falling edges of the DAC output. Hence, clock jitter errors are directly added to the input signal [31]. Another important limitation of CT-BPΣΔMs is the *excess loop delay* contributed by each building block in the modulator loop, which can severely degrade the quantization noise transfer function [32].^{†6}

Fig. 11.19(a) is a conceptual block diagram of a CT-BPΣΔM. This modulator is internally a DT circuit since there is an Sampling-and-Hold (S/H) circuit inside the loop, just at the quantizer input. This fact makes the overall loop from the output of the quantizer back to its input have a *Z*-domain transfer function as illustrated in Fig. 11.19(b). The equivalent DT loop filter transfer function is [22]:

$$\begin{aligned} H(z) &= Z\left[L^{-1}[DAC(s)H(s)]\Big|_{t=nT_s}\right] = \\ &= Z\left[\int_{-\infty}^{\infty} DAC(\tau)h(t-\tau)d\tau\Big|_{t=nT_s}\right] \end{aligned} \quad (11.18)$$

where $DAC(t)$ is the impulsive response of the DAC.

The expression in eq.(11.18), known as *pulse invariant transformation*, allows us to obtain an equivalent relation between DT- and CT-BPΣΔMs. Thus, the synthesis process of a CT-BPΣΔM starts from a DT loop filter that satisfies the required specifications and then it is transformed into an equivalent CT filter

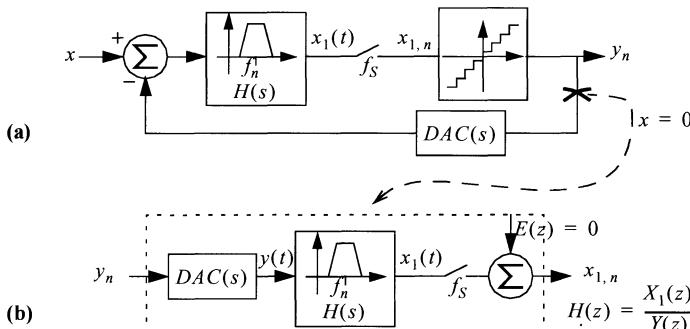


Figure 11.19: Basic architecture of a CT-BPΣΔM. (a) Conceptual block diagram.
(b) Open loop block diagram.

6 . Section 11.5 will treat these limitations in more detail.

using eq.(11.18). Therefore, much of the knowledge available for DT-BPΣΔMs can be utilized for synthesizing CT-BPΣΔM architectures.

There are different ways of realizing the DT-to-CT transformation in eq.(11.18) depending on the shape of $DAC(t)$, namely: NonReturn-to-Zero (NRZ), Return-to-Zero (RZ) and Half-delay Return-to-Zero (HRZ). The differences among them will originate several architecture issues, specific of CT-BPΣΔMs. A detailed analysis of those issues – beyond the scope of this Chapter – can be found in several works related to this subject [20][22].

11.5 BUILDING BLOCKS AND ERROR MECHANISMS IN BPΣΔMs

The BPΣΔM architectures described in previous sections have been considered ideal except for the quantization error. In practice, the behaviour of such architectures deviates from the ideal performance as a consequence of their building block error mechanisms. This section discusses the impact of circuit parasitics on the performance of BPΣΔMs, showing their effects on the N_{TF} , the in-band noise power and the harmonic distortion.

The study presented here will focus on a single-loop 4th-order BPΣΔM (4th-BPΣΔM) – derived from applying the transformation in eq.(11.12) to a 2nd-order LPΣΔM. These modulators are easy to understand and simple to design, are capable of providing high resolution together with large tolerance to imperfections and robust stable operation [12]. Nevertheless, this study can be easily extended to other architectures such as multi-stage cascade architectures [33]. In these architectures, the error contributions due to the first stage – usually a single-loop BPΣΔM like that treated in this section – constitute the most significant degrading factor of the overall modulator performance.

As a starting point for our study, the resonator – the main block of BPΣΔMs – is analyzed. Several architectures are described as well as their circuit implementation using different circuit techniques.

11.5.1 From integrators to resonators

As stated in Section 11.4.1, most of the reported BPΣΔM architectures have been obtained from corresponding lowpass prototypes by applying the transformation in eq.(11.12). As a consequence of this transformation, the integrators which form the loop filter in the original modulator become resonators with the following transfer function:

$$H_{res}(z) = \frac{\pm z^{-a}}{1 + z^{-2}} \quad (0 < a \leq 2) \quad (11.19)$$

which has their poles located at $z_n = \exp(\pm 2\pi j)$, that is, $f_n = f_S/4$.

There are many filter structures which implement the transfer function in eq.(11.19). Fig. 11.20 shows three alternatives which have been used in BPΣΔMs. Fig. 11.20(a) is based on two delay elements connected in a loop, and is often called Delay-loop resonator [34]. Fig. 11.20(a)-(b) are based on Lossless Direct Integrators (LDIs) and Forward-Euler Integrators (FEIs), referred to as LDI-loop and FE-loop resonators, respectively [14]. Assuming that the scaling coefficients are $A_F = 1$ and $A_{FB,FB2} = -2$, the three resonators in Fig. 11.20 are identical. They have the transfer function in eq.(11.19) with $a = 2$ for Fig. 11.20(a) and (c), and $a = 1$ for Fig. 11.20(b).

In the presence of errors, the scaling coefficients A_F and $A_{FB,FB2}$ deviate from their nominal values due to either capacitor ratio errors – for SC circuits [15] – or to transistor size ratio errors – in the case of SI circuits [18]. As a result of these errors, the poles of $H_{res}(z)$ experience movements around their nominal

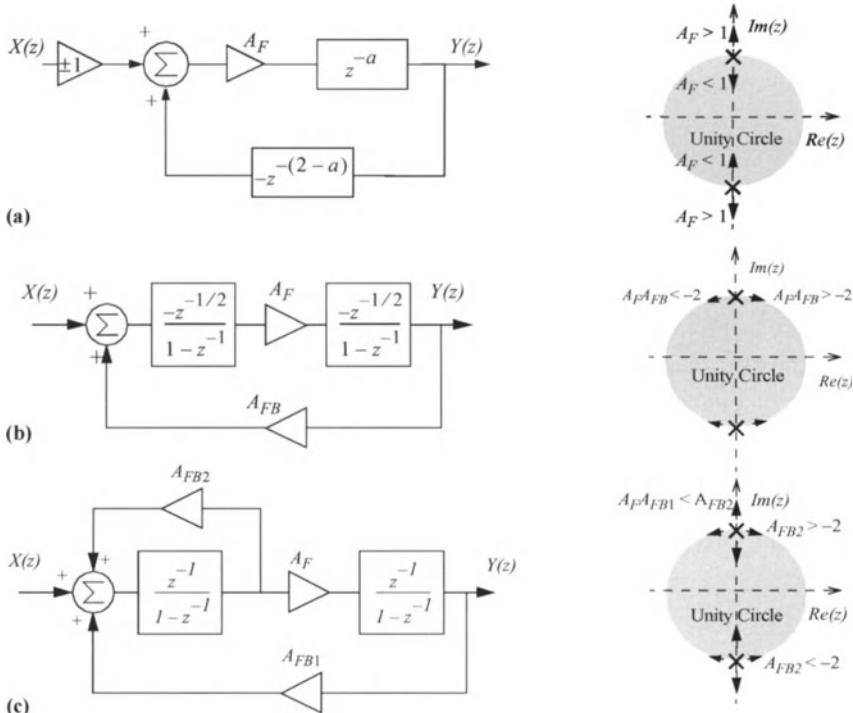


Figure 11.20: Different filter implementations of the resonator transfer function and movement of their poles under errors in their feedback gains. a) Delay-loop. b) LDI-loop. c) FE-loop.

positions – different for each resonator structure as shown in Fig. 11.20.

In some structures the filter poles move around the unity circle in the Z -plane. This results in the resonant frequency, f_n , not being properly placed. This is the case of LDI-loop resonators and FE-loop resonators under changes on $A_{FB}, FB2$. In the Delay-loop and FE-loop structures, the effect of errors will move the resonator poles off the unit circle, causing instability. If the poles move inside the unit circle, then the Q factor will be reduced, thus reducing the gain of H_{res} at f_n .

Although the possible instability of FE-loop resonators appears to be a drawback as compared to LDI-loop resonators, some authors propose to design filters with a small instability with the objective of reducing idle tones in $\Sigma\Delta$ Ms [35]. This is still a matter of discussion. In fact, the authors in [14] reported similar experimental results from two 2nd-order BP $\Sigma\Delta$ Ms, one of them based on LDI-loop resonators and the other one using FE-loop resonators.

The resonator architectures shown in Fig. 11.20 can be implemented using DT circuit techniques. As an illustration, Fig. 11.21 shows the schematics of the resonators in Fig. 11.20 using SC Fully Differential (FD) circuits^{†7} [14][34], and Fig. 11.22 shows the corresponding SI (single-ended) realizations[18]. All these resonators have their poles placed at $f_n = f_S/4$. This is a consequence of the $z^{-1} \rightarrow -z^{-2}$ transformation. However, in some applications such as multi-standard radio receivers, it could be interesting to design the scaling coefficients to be programmable. This will allow us to control f_n without changing f_S . This can be

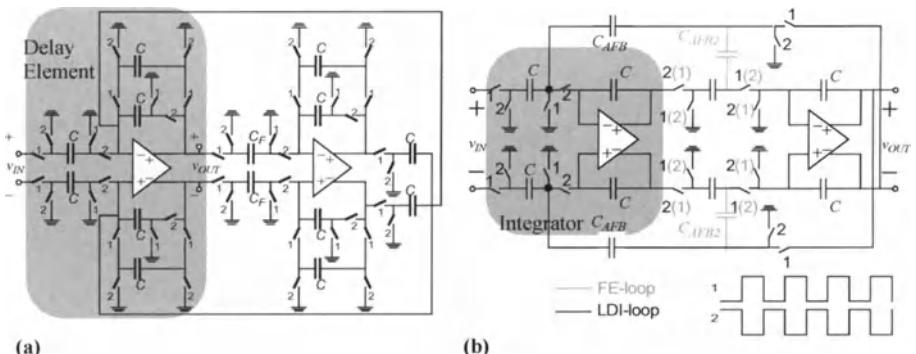


Figure 11.21: SC FD realizations of the resonators in Fig. 11.20. a) Delay-loop.
b) LDI and FE-loop.

7. SC resonators can be also implemented using a two-path architecture [33][35] – not shown in this Chapter for the sake of simplicity. Their main advantage is that only one opamp is required instead of two as in Fig. 11.21. However, a complex clock phase scheme is used, which needs to be carefully timed in order to avoid image components to appear in the signal band.

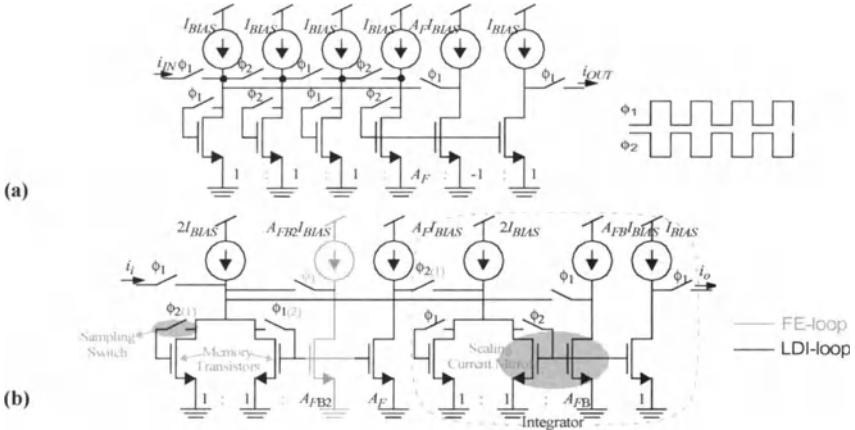


Figure 11.22: SI realizations of the resonators in Fig. 11.20. a) Delay-loop and clock phase generator. b) LDI and FE-loop.

done, for instance, by changing $C_{AFB,AFB2}/C$ in Fig. 11.21 and the current mirror output transistor size, $A_{FB,FB2}$, in Fig. 11.22(b).

In case of CT ($g_m - C$) resonators like that shown in Fig. 11.23, the resonant frequency, given by:

$$f_n = \frac{1}{2\pi} \sqrt{\frac{g_{m2}g_{mR}}{C_1 C_2}} \quad (11.20)$$

can be varied by making g_{mR} to be programmable electronically [37].

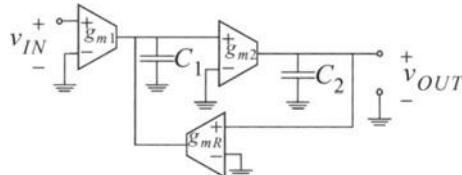


Figure 11.23: Conceptual schematic of a $g_m - C$ resonator.

11.5.2 Quantization noise shaping degradation due to circuit errors

As discussed in previous section, LDI-loop resonators are the only ones which remain stable under changes in their scaling coefficients. For that practical reason, this has been the resonator structure chosen for our study^{†8}. Fig. 11.24 shows the

⁸ . A similar discussion to that presented here can be followed for the case of either FE-loop or Delay-loop based BPΣΔMs.

Z-domain block diagram of a 1-bit 4th-BPΣΔM based on LDI-loop resonators. Note that the required feedback loop delay has been realized through two additional delay blocks. The scaling factors can be optimized to obtain similar dynamic range for both resonators. This yields to the following nominal values,

$$A_{RES2} = A_{DAC2} = -A_{DAC1} = 1 \quad A_{RES1} = 1/2 \quad (11.21)$$

Ideally, the Z-transform of the modulator output is given by eq.(11.2) with

$$\begin{aligned} S_{TF} &= z^{-2} \\ N_{TF} &= (1 + z^{-2})^2 \end{aligned} \quad (11.22)$$

and P_Q and SNR are respectively given by eq.(11.8) and eq.(11.13); with $f_n = f_S'/4$, $N = 1$ and $L = 2$. However, such an ideal performance can only be achieved provided that the resonators in Fig. 11.24 are realized without errors.

In the case of SC realizations, the major sources of error that degrade the noise-shaping of BPΣΔMs are [15]:

- *Finite operational amplifier (opamp) DC gain*, represented by A_V .
- *Incomplete settling error*, ε_s , caused by the limited opamp bandwidth. In a first-order approach (single-pole), $\varepsilon_s \equiv \exp[-T_S'/(2\tau)]$, where τ is the closed-loop time constant of the SC integrator.
- *Mismatch capacitor ratio error*, ε_m , in the scaling coefficients (C_{AFB}'/C in Fig. 11.21(b)).

In the case of SI BPΣΔMs, the main circuit parasitics are [18]:

- *Finite conductance ratio error*, defined as $\varepsilon_g \equiv 2(g_o/g_i)$, where g_o and g_i are respectively the output and input conductances of SI memory cells.
- *Charge injection error*, ε_q , due to the charge injected, δq , by the sampling switch onto the storing capacitance, C_{gs} , of the memory transistor (see Fig. 11.22(b)).
- *Incomplete settling error*, ε_s , caused by the finite bandwidth, $g_m'C_{gs}$, where g_m is the transconductance of the memory cell.

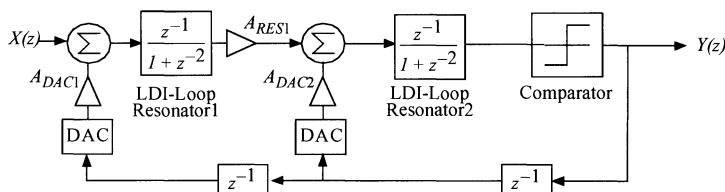


Figure 11.24: LDI-loop based 4th-order 1-bit BPΣΔM architecture under study.

- *Mismatch error*, ϵ_m , in β and V_T , of the current mirror transistors used to implement the scaling coefficients (A_{FB} in Fig. 11.22(b)).

In the presence of the above-mentioned errors, the resonator transfer function is modified into [15][18]

$$H_{res}(z) \equiv \frac{(1-\mu)z^{-1}}{1 + \xi_1 z^{-1} + (1 - \xi_2)z^{-2}} \quad (11.23)$$

where μ , ξ_1 and ξ_2 are different for each error, as Table 11.2 shows.

Replacing the transfer functions of the resonators in Fig. 11.24 with eq.(11.23), the erroneous N_{TF} of the 4th-BPΣΔM is obtained, giving:

$$N_{TF}(z) \equiv [1 + \xi_1 z^{-1} + (1 - \xi_2)z^{-2}]^2 \quad (11.24)$$

The zeroes of N_{TF} are shifted from their nominal positions at $f_S/4$, thus degrading the filtering performed by the resonators and making the quantization noise in-band power, and correspondingly, the SNR to decrease.

From eq.(11.7)-eq.(11.10) and eq.(11.24), it can be shown that the erroneous P_Q and DR are approximately given by [18]:

Table 11.2: Resonator transfer function degradation with circuit errors.

SC CIRCUITS	μ	ξ_1	ξ_2
A_V	$\frac{2}{A_V} \left(1 + \frac{C_S}{C_I}\right)$	$\frac{-2}{A_V} \left(2 + \frac{C_S}{C_I}\right)$	$\frac{2}{A_V} \frac{C_S}{C_I} \dagger *$
$\epsilon_s \equiv \exp\left(\frac{-T_S}{2\tau}\right)$	$2\epsilon_s$	$-4\epsilon_s$	0
$\epsilon_m \equiv \frac{\delta C_{AFB}}{C_{AFB}} - \frac{\delta C}{\delta C}$	ϵ_m	ϵ_m	0
SI CIRCUITS			
$\epsilon_g \equiv \frac{2g_o}{g_i}, \epsilon_q \equiv \frac{\delta q}{C_{gs}(V_{GS} - V_T) _Q}$	$2\epsilon_{g,q}$	0	$4\epsilon_{g,q}$
$\epsilon_s \equiv \exp\left(\frac{-T_s g_m}{2C_{gs}}\right)$	$2\epsilon_s$	$-4\epsilon_s$	$4\epsilon_s$
$\epsilon_m \equiv \frac{\delta \beta}{\beta} - \frac{\delta V_T}{(V_{GS} - V_T) _Q}$	ϵ_m	ϵ_m	0

* C_s and C_I represent the sampling and the integration capacitances, respectively.

$$P_Q = \frac{\pi^4 \Delta^2}{60M^5} (\nabla P_Q) \quad DR \equiv \frac{15M^5}{2\pi^4 (\nabla P_Q)} \quad (11.25)$$

where

$$\nabla P_Q \equiv 1 + \frac{10}{3}(3\xi_1^2 + \xi_2^2)\left(\frac{M}{\pi}\right)^2 + 5(\xi_1^2 + \xi_2^2)^2\left(\frac{M}{\pi}\right)^4 \quad (11.26)$$

is the in-band quantization noise increase caused by circuit parasitics, which in combination with the expressions for ξ_1, ξ_2 given in Table 11.2, allow us to know the maximum error permitted as a function of M for a given DR .

In addition to the DR loss, circuit parasitics cause a shifting of the position of f_n , denoted as δf_n . Solving the roots of eq.(11.24) and assuming that $\xi_1, \xi_2 \ll 1$, it can be shown that:

$$\delta f_n \equiv f_n - \frac{f_s}{4} \equiv \xi_1 \frac{M}{\pi} \left(\frac{B_w}{2} \right) \quad (11.27)$$

Thus, depending on the way that the noise shaping of BPΣΔMs is degraded by circuit errors, they can be grouped in the following families:

- Those errors that cause $\xi_1 = 0, \xi_2 \neq 0$, and hence, their main effect is to reduce the Q-factor of the resonator transfer function, thus lowering the bandstop attenuation of the modulator bandpass filtering, increasing P_Q as illustrated in Fig.11.25(a).
- Those errors in which $\xi_1 \neq 0, \xi_2 = 0$, whose main effect is to change f_n . However, P_Q does not significantly increase as shown in Fig.11.25(b).
- Those errors in which $\xi_1 \neq 0, \xi_2 \neq 0$, causing a combined effect of increasing P_Q and shifting f_n , as illustrated in Fig.11.25(c).

All these results have been validated by time-domain behavioural simulation

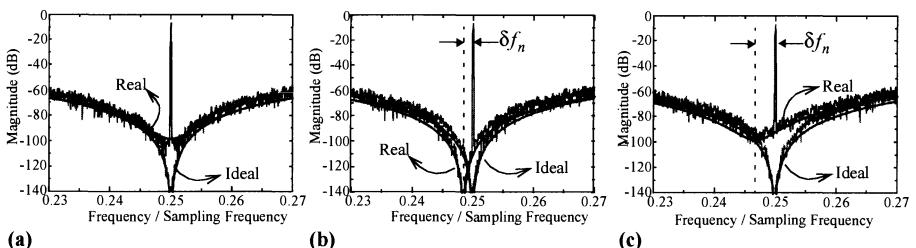


Figure 11.25: Noise-shaping degradation with circuit errors: a) $\xi_1 = 0, \xi_2 \neq 0$. b) $\xi_1 \neq 0, \xi_2 = 0$. c) $\xi_1 \neq 0, \xi_2 \neq 0$.

[15][18]. Thus, Fig.11.25 compares the theoretical model – solid line – with simulation showing a good agreement between simulation and theory.

11.5.3 Harmonic distortion in bandpass $\Sigma\Delta$ modulators

In Section 11.5.2, circuit errors have been assumed to be linear. However, practically all errors are signal-dependent and hence, in addition to degrade N_{TF} , cause Harmonic Distortion (HD) in BP $\Sigma\Delta$ Ms.

As an illustration, Fig. 11.26 shows the increase of the 3rd-order InterModulation distortion^{†9}, IM_3 , due to non-linear settling in FD 4th-BP $\Sigma\Delta$ Ms based on SI circuits. In this case, the major source of non linearity is the signal-dependent, g_m , which causes main SI errors (ε_g , ε_q and ε_s) to be non-linear [18].

As in LP $\Sigma\Delta$ Ms, the non-linearity of voltage-mode (SC and $g_m - C$) BP $\Sigma\Delta$ Ms is mainly caused by the following error mechanisms [25]:

- *Slew-Rate*, due to the saturation of the opamp in current, causing a non-linear transient response of the integrator.
- *Non-linear open-loop DC-gain of the opamp*, due to the fact that the transition between the linear and saturation output region is gradual. This effect can be modelled as polynomial dependence of A_V on the opamp output (input), $A_V = A_0(1 + \gamma_1 v + \gamma_2 v^2 + \dots)$.^{†10}
- *Non-linear capacitors*, due to the dependency of the value of the capacitance on its stored voltage; modeled as $C(v) = C^o(1 + \alpha v + \beta v^2 + \dots)$, where α, β, \dots are non-linear technology-dependent coefficients.

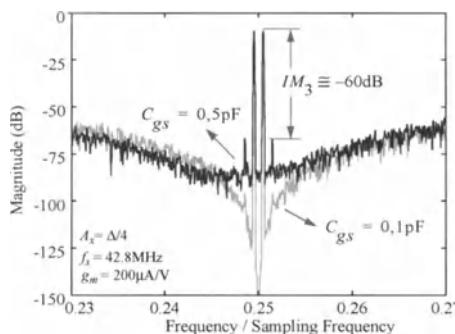


Figure 11.26: Harmonic Distortion in SI BP $\Sigma\Delta$ Ms due to non-linear settling.

- 9 . In bandpass signal processing, IM_3 is a more appropriate figure than HD_3 for measuring HD. Assuming an input signal formed by two tones at f_2, f_1 , IM_3 is defined as the amplitude of the output at $2f_2 - f_1$ and $2f_1 - f_2$ related to the output amplitudes at f_2, f_1 .
- 10 . In CT-BP $\Sigma\Delta$ Ms, the effect caused by the non-linear transconductor is similar to that due to *SR* and the non-linear A_V in SC circuits [20].

Considering the effect of non-linear errors, the behaviour of a LDI-loop resonator can be generically expressed by the following equation [18]^{†11}:

$$\begin{aligned} x_{o,n} = & (1 - \mu)x_{i,n-1} - \xi_1 x_{o,n-1} - (1 - \xi_2)x_{o,n-2} \\ & + F_{NL}(x_{i,n}, x_{i,n-1}, \dots, x_{o,n}, x_{o,n-1}, \dots, \overline{\xi_{NL}}) \end{aligned} \quad (11.28)$$

where F_{NL} represents the non-linearity, which is function of the resonator input, x_i , the output, x_o , and a vector of electrical parameters, $\overline{\xi_{NL}}$ – different for each error mechanism.

The obtainment of design equations for IM_3 requires solving the time-domain equations that govern the behaviour of BPΣΔMs, considering that the first resonator^{†12} behaviour is given by eq.(11.28).

Instead of deriving cumbersome expressions of IM_3 for each signal-dependent parasitics – beyond the scope of this Chapter^{†13} –, we will center our attention on a critical effect: the non-linear sampling process. Contrary to the lowpass case, this error mechanism constitutes one of the most limiting factors in BPΣΔMs used in modern telecommunication systems.

In the front-end of such systems, the input (analog) signal is changing very quickly during the sampling phase interval. As an illustration, Fig. 11.27 shows the transient evolution of a sinusoidal signal of amplitude X and frequency $f_i \equiv f_S/4$, when sampled at f_S . In this case, corresponding to a typical input sig-

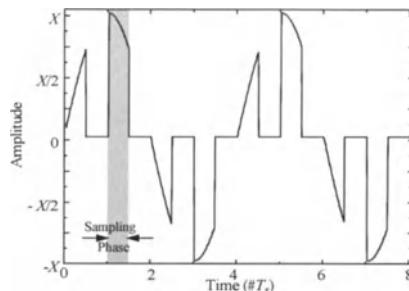


Figure 11.27: Transient evolution of a input sinusoidal signal of amplitude X an frequency $f_i \equiv f_S/4$ when sampled at $f_S/4$.

11 .In this Chapter, x_n denotes $x(nT_S)$.

12 .The contribution of the second resonator to HD is attenuated by the gain of the first resonator in the signal band. For this reason, only the first resonator contribution has to be considered.

13 .The interested reader is referred to [18] where a complete analysis of HD in SI BPΣΔMs is described for each non-linear error.

nal in BPΣΔMs, the signal amplitude can change up to $X/\sqrt{2}$ during the sampling phase. This change leads to a non-linear transient response of the sampling circuit which manifests as HD at the output of the BPΣΔM.

For a better understanding of this phenomenon, let us consider the input stage of the SC resonators shown in Fig. 11.21(b). During the clock phase 1 (sampling phase), the input signal, v_{IN} , is stored in the capacitor C . In practice, this circuit is realized by using CMOS switches as shown in Fig. 11.28(a). These switches are sized such that the *switch-on* resistance, R_{on} , verifies $R_{on}Cf_s \ll 1$, thus making the incomplete settling error negligible. This condition is not difficult to satisfy in practice. However, R_{on} strongly depends on the input signal amplitude as illustrated in Fig. 11.28(b) for a CMOS switch in a 2.5V-0.25μm CMOS technology. As a consequence of this non-linearity, the sampling circuit will cause HD. Based on the Volterra series method, the authors in [38] demonstrated that, in FD SC circuits like that in Fig. 11.28(a), the HD due non-linear sampling is given by:

$$HD_3 \equiv \frac{\pi f_i C R_{on}}{2(V_{ON} - V_T)^2} V_{IN}^2 \quad (11.29)$$

where V_{ON} stands for the switch-on voltage^{†14}.

Note that, HD_3 increases with the input frequency. This is because, as the input signal frequency becomes higher, there will be a larger variation of the signal amplitude during the sampling phase, thus enlarging the effect of the R_{on} non-linearity.

An important conclusion derived from eq.(11.29) is that, even for low values of the incomplete settling error, high levels of HD can be obtained. This suggests

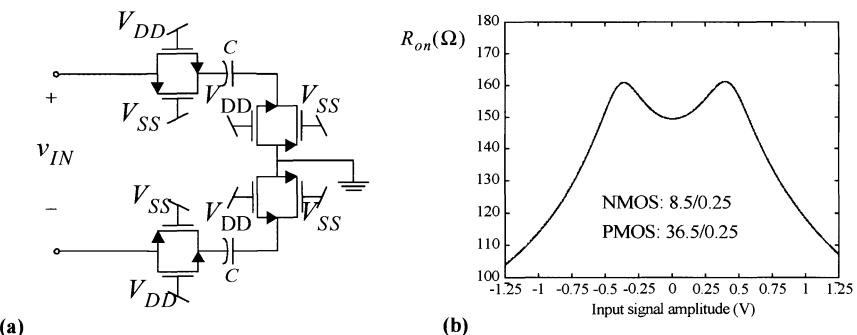


Figure 11.28: Non-linear sampling in SC BPΣΔMs. a) Sampling circuit. b) DC characteristic of CMOS switches (2.5V-0.25μm CMOS technology).

14 .The study developed in [38] is based on NMOS switches instead of CMOS switches. In this case, $V_{ON} = V_{DD}$.

that, in some applications, either clock-boosting or similar techniques should be used to reduce the effect of non-linear sampling [36].

A similar conclusion can be found in SI BP $\Sigma\Delta$ Ms [18]. In this case, the HD due to dynamic errors is dominated by the non-linear sampling of the front-end memory cell. As an illustration, Fig. 11.29 compares the HD caused by the non-linear sampling with that caused by the incomplete settling. Note that, in practical cases, that is, $\epsilon_s < 0.1\%$, the HD is dominated by the non-linear sampling.

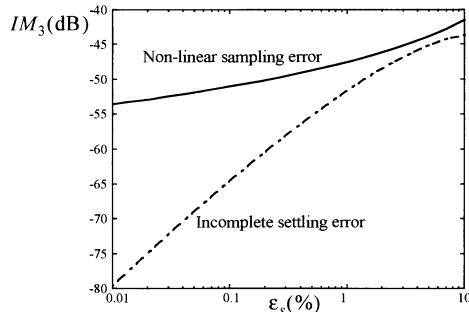


Figure 11.29: Comparison of the HD caused by the incomplete settling error and the non-linear sampling in SI BP $\Sigma\Delta$ Ms.

11.5.4 Thermal noise

Thermal noise constitutes the ultimate limiting factor of BP $\Sigma\Delta$ Ms. As for the LP $\Sigma\Delta$ M case, only the thermal noise contributions at the input node of the modulator will be important because they are added directly to the input signal, thus appearing with no filtering in the output spectrum.

In DT BP $\Sigma\Delta$ Ms, as a consequence of the S/H operation, the input-referred noise PSD increases with respect to the CT case. This increase is proportional to the relationship between the equivalent noise bandwidth and the sampling frequency. For that reason, we will center on DT (SC or SI) realizations.

Fig. 11.30(a) shows the most important thermal noise sources in a LDI-based 4th-BP $\Sigma\Delta$ M. In this block diagram, v_{nm} represents the input-equivalent noise source of the m -th integrator. In practical cases, noise contributions of DACs can be considered negligible as compared to that due to resonators. In addition, the noise contributions of both the 3rd- and the 4th- integrators are attenuated by the gain of the 1st resonator in the signal bandwidth. However, the contribution of the 2nd integrator, v_{n2} , has to be considered as illustrated in the output spectrum of Fig. 11.30(b).

It can be shown that the contribution of v_{n2} to the input-equivalent noise, v_{ni} , is twice that of that due to v_{n1} because in the signal band [18],

$$\left| 1 - z^{-1} \right|^2 \Big|_{z = e^{-j2\pi B_w T_s}} \cong 2 \quad (11.30)$$

Therefore, the in-band power of v_{ni} is approximately given by:

$$P_{ni} \cong P_{n1} + 2P_{n2} \quad (11.31)$$

where P_{n1} and P_{n2} are respectively the in-band power of v_{n1} and v_{n2} .

In case that $P_{ni} > P_Q$, the *SNR* and the *DR* for a sinusoidal input signal of amplitude $A = \alpha\Delta/2$ are respectively given by:

$$SNR_{th} \cong \frac{\alpha^2 \Delta^2}{8P_{ni}} \quad DR_{th} \cong \frac{\Delta^2}{8P_{ni}} \quad (11.32)$$

11.5.5 Jitter noise

In previous sections, ideal clock phase signals have been considered. In practice, the period of the clock signal presents random variations in its nominal value as illustrated in Fig. 11.31(a). This is due to certain intrinsic uncertainties in the time in which clock transitions occur, known as *jitter*. The result is a non-uniform sampling, responsible for extra (white) noise at the output of $\Sigma\Delta$ Ms [12].

As a consequence of jitter, it can be shown that the *SNR* of DT-BP $\Sigma\Delta$ Ms is degraded as [31]:

$$SNR|_{DTj} = \frac{M}{4\pi^2 \sigma_j^2 f_n^2} \quad (11.33)$$

where σ_j is the standard deviation of the sampling time error, δt_j ¹⁵.

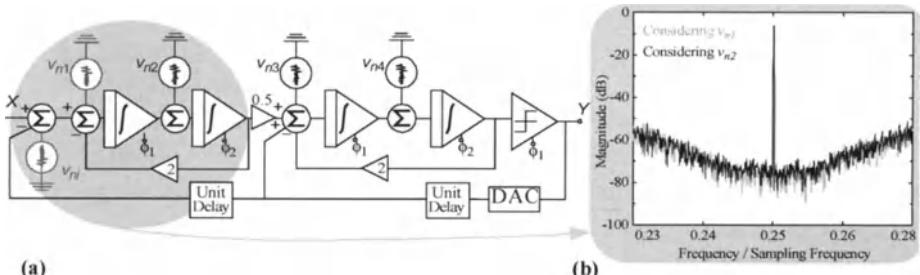


Figure 11.30: Thermal noise in LDI-loop based 4th-BP $\Sigma\Delta$ Ms. a) Most important noise sources. b) Contribution to the output spectrum noise.

15 .The clock jitter error, represented here by parameter δt_j , behaves as a random variable with a Gaussian distribution of standard deviation σ_j and zero mean.

Note that, as in LP $\Sigma\Delta$ Ms, $SNR|_{DTj}$ decreases with the signal frequency. However, in BP $\Sigma\Delta$ Ms the signal frequency is a substantial fraction of f_S , typically $f_n \approx f_S/4$, what constitutes a more limiting factor. As an illustration, Fig. 11.31(b) shows the increase of the in-band jitter noise with f_S in BP $\Sigma\Delta$ Ms.

Jitter noise is specially critical in CT BP $\Sigma\Delta$ Ms. In this type of BP $\Sigma\Delta$ Ms, clock jitter comes from two sources^{†16}: the S/H circuit and the DAC. The former is subject to the same filtering than the quantization error, and hence, will have a small effect on the modulator performance. However, the DAC jitter noise is directly added with the input signal, thus increasing the in-band noise power. This power increase will depend on the impulsive response of the DAC. Thus, as illustrated in Fig. 11.31(c), for the same bit sequence, the number of DAC output rising/falling edges per clock cycle will depend on the type of DAC response. This will manifest as a different SNR degradation, given by [31]:

$$SNR|_{CTj} \cong \begin{cases} \frac{\text{sinc}(\pi f_n T_S)}{64 \sigma_j^2 B_w^2 M} & \text{for NRZ DACs} \\ \frac{\text{sinc}(\pi f_n T_S)}{64 \left(\frac{T_S}{T_0}\right)^2 \sigma_j^2 B_w^2 M} & \text{for RZ DACs} \end{cases} \quad (11.34)$$

Note that, as Fig. 11.31(c) predicts, a lower SNR degradation is achieved by using NRZ DACs. Thus, comparing eq.(11.33) and eq.(11.34) it can be shown

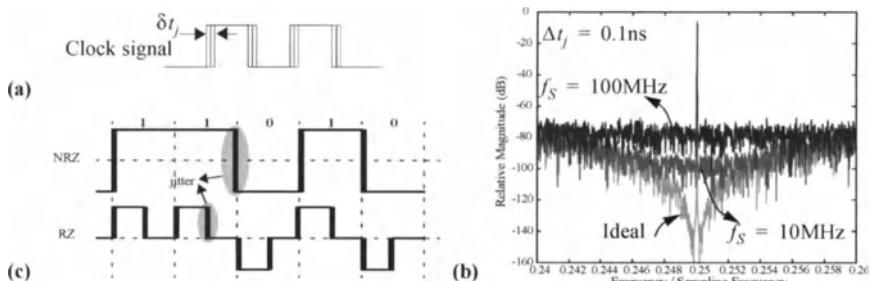


Figure 11.31: Jitter noise. a) Uncertainties in the clock transitions. b) Effect on DT-BP $\Sigma\Delta$ Ms. c) Jitter in both NRZ and RZ DAC output responses.

16 .There is another source of jitter noise caused by comparator *metastability*. This non-ideal effect causes an increase of the in-band noise which can be modelled as a jitter noise as shown in [20]. This kind of jitter is often referred to as signal-dependent jitter because it is due to the signal-dependent comparator delay.

that CT-BP $\Sigma\Delta$ M s are more sensitive to jitter than DT-BP $\Sigma\Delta$ M s . For instance, considering NRZ DACs in eq.(11.34) (the best case) and $f_n \equiv f_S/4$, the condition

$$\sigma_j^2|_{CT} = \sigma_j^2|_{DT} \left(\frac{\pi}{8}\right)^2 \text{sinc}\left(\frac{\pi}{4}\right) \approx 0.14 \sigma_j^2|_{DT} \quad (11.35)$$

should be satisfied to obtain the same *SNR* loss in both CT- and DT-BP $\Sigma\Delta$ M s .

Besides jitter, the DAC time response causes another important degradation in CT-BP $\Sigma\Delta$ M s which is analyzed in next section.

11.5.6 Excess loop delay in continuous-time BP $\Sigma\Delta$ M s

Ideally, the DAC output currents^{†17} should respond immediately to the quantizer clock edge. In practice, there exists a delay, as illustrated in the *multi-feedback* CT-BP $\Sigma\Delta$ M^{†18} of Fig. 11.32(a). This delay, often referred to as *excess loop delay* or *loop delay*, is normally expressed as a fraction of the sampling period, $\tau_d = \rho_d T_s$, where $0 < \rho_d \leq 1$.

As a consequence of the poles introduced by $\rho_d \neq 0$, the noise-shaping transfer functions, S_{TF} and N_{TF} , are modified, thus increasing the in-band noise power as illustrated in Fig. 11.32(c). This effect becomes specially critical in telecom applications, in which large values of $\rho_d - \rho_d \approx 0.4$ in the example of Fig. 11.32(b)-(c) – may eventually make CT-BP $\Sigma\Delta$ M s unstable [20].

Mathematically speaking, complex analyses are required to obtain the stability condition in most common BP-CT $\Sigma\Delta$ M architectures. Instead of that, simulation-based analyses are normally used to know the critical value of ρ_d leading to instability. This simulation-based approach is used – in combination with other strategies such as coefficient tunning and adding extra DAC branches – to compensate for the excess loop delay as stated in [20].

11.6 STATE-OF-THE-ART BANDPASS $\Sigma\Delta$ ADC s

Although the idea of translating the zeroes of N_{TF} from DC to any given IF location was originally presented by Schreier and Snelgrove in 1989 [5], it was

17 .In most of CT-BP $\Sigma\Delta$ M s , DACs are realized by differential pairs driven by the quantizer output.

18 .This architecture uses different types of DACs – each with separately feedback coefficients – to obtain similar noise-shaping than that of the DT-BP $\Sigma\Delta$ M in Fig. 11.24 [32]. This strategy allows us to use resonator transfer function of the type $s/(s^2 + \omega^2)$, which are easier to realize through $g_m C$ circuits than the CT filters resulting from a DT-to-CT transformation like that in eq.(11.18).

not until 1992 that Jantzi, Snelgrove and Ferguson published the first monolithic BPΣΔM IC [39]. Since then, there have been a large number of ICs implemented in several technologies, with diverse circuit techniques; using different supply voltages.

Table 12.3 shows a summary of the CMOS BPΣΔ ADC ICs published to this day. For each of them, the most significant figures are shown, namely: DR , SNR , f_S , f_n , B_w , the power consumption, the characteristics of the fabrication process and the modulator architecture. For the latter, the synthesis method employed is also given. Note that most modulators use a single-loop architecture, obtained by applying the $z^{-1} \rightarrow -z^{-2}$ transformation, which, as stated in Section 11.4.1, makes $f_n = f_S/4$.

The BPΣΔMs in Table 11.3 cover multiple applications in digital wireless communications, ranging from telemetering [40] to digital radio receivers [13] [14][19][23][33] and modern cellular phones [22][34][41][42]. Depending on the application, there are different specifications for DR , B_w and f_n . Hence, it is difficult to state a measure of comparison among all reported modulators. Thus, in

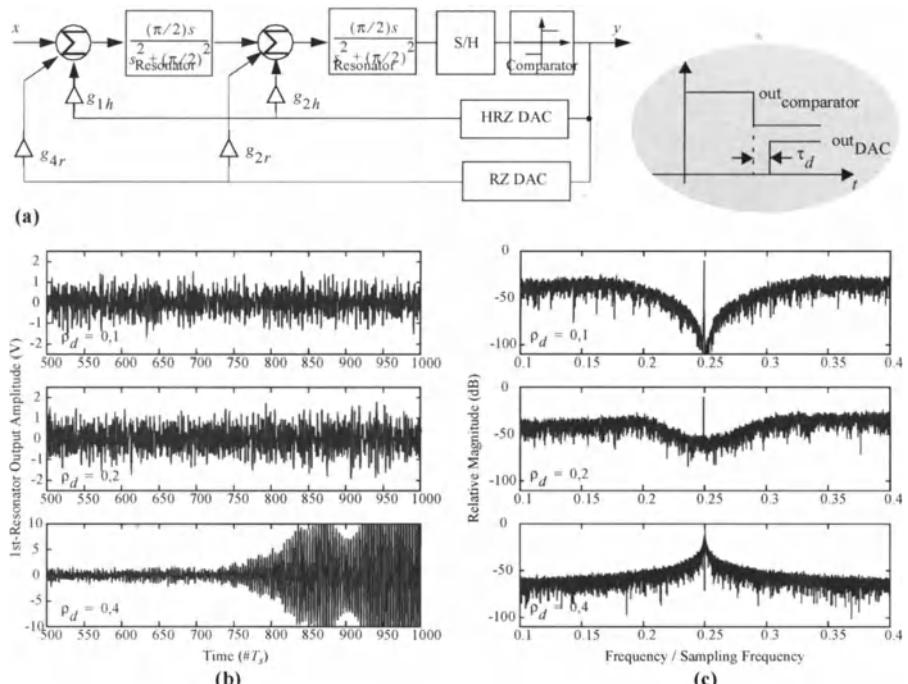


Figure 11.32: Excess loop delay. a) DAC time response in CT-BPΣΔMs. Effect on: b) Transient response of first resonator. c) Modulator output spectrum.

Table 11.3: Summary of CMOS bandpass $\Sigma\Delta$ modulators published up to now.

Author ^{†*}	DR (bit)	f_S (MHz)	f_n (MHz)	B_w (kHz)	Power (mW)	Technology	Architecture
Jantzi'93 [13]	10.2	1.825	0.455	8	210	3 μ m DP / 5V	4th, Optim. N_{TF}
Longo'93 [34]	15	7.2	1.8	30	--	1 μ m DP / 5V	4th, $z^{-1} \rightarrow -z^2$
Song'95 [41]	9	8	2	30	0.8	2 μ m DP / 3.3V	4th, $z^{-1} \rightarrow -z^2$
Hairapetian'96 [33]	11.7	13	3.25	200	14.4	0.8 μ m DP / 3V	4-2, $z^{-1} \rightarrow -z^2$
André'96 [43]	8	8	2	64	8	0.5 μ m DP / 3.3V	6th, $z^{-1} \rightarrow -z^2$
Liu'97 [44]	11.8	0.827	0.413	2	--	2 μ m DP / 5V	4th, Optim. N_{TF}
Cormier'97 [21]	9.5	1.25	0.25- 0.375	6.25	--	2 μ m DP / 5V	4th, $z^{-1} \rightarrow -z^2$
Ong'97 [30]	12.2	80	20	200	72	0.6 μ m SP / 3.3V	4th, $z^{-1} \rightarrow -z^2$
Jantzi'97 [28]	10.8	10	3.75	200	130	0.8 μ m SP / 5V	4th, Quadrature
André'98 [45]	9.2	4	1	200	19	0.5 μ m DP / 3V	3th-3bit, Optim. N_{TF}
Bazarjani'98 [46]	6.7	40	20	1250	65	0.5 μ m DP / 5V	2nd, $z^{-1} \rightarrow -z^2$
Chuang'98 [47]	13	0.5	0.125	0.5	--	2 μ m DP / 5V	6th, Optim. N_{TF}
Van Engelen'99 [19]	11.7	40	9.15	200	60	0.5 μ m DP / 5V	6th, CT
	10.8	80	10.7	200			
Park'99 [48]	12.2	20	5	200	180	0.65 μ m SP / 4V	4th, $z^{-1} \rightarrow -z^2$
Tabatabaei'99 [42]	13	80	20	1250	90	0.25 μ m DP / 2.5V	6th, $z^{-1} \rightarrow -z^2$
Tonietto'99 [49]	12.7	42.8	10.7	200	80	0.35 μ m SP / 3.3V	6th, Optim. N_{TF}
Bazarjani'99 [50]	9.4	68	17	1250	48	0.6 μ m DP / 3V	4-4, $z^{-1} \rightarrow -z^2$
Tao'99 [51]	8	400	100	200	330	0.35 μ m SP / 3.3V	2nd, CT
Tabatabaei'00 [52]	12	64	16	2000	110	0.25 μ m SP / 2.5V	6th, 2-path
Cusitano'00 [53]	12	37.05	10.7	200	116	0.35 μ m SP / 3.3V	6th, Optim. N_{TF}
Cheung'01 [54]	6.7	42.8	10.7	200	12	0.35 μ m SP / 1V	2nd, $z^{-1} \rightarrow -z^2$
Salo'02 [36]	11.7	80	20	270	56	0.35 μ m?P / 3V	4th, $z^{-1} \rightarrow -z^2$
	6.7			3840			
Ueno'02 [55]	12.6	10	0.566	250	77	0.25 μ m DP / 2.5V	2-2, 3bit

* Sorted by date of publication.

order to compare different performances, the following *Figure-Of-Merit (FOM)* will be used for BP $\Sigma\Delta$ Ms^{†19}:

$$FOM_{BP} = \frac{\text{Power (mW)}}{2^{DR(\text{bits})} \times f_n(\text{MHz})} \quad (11.36)$$

Fig. 11.33, plots FOM_{BP} vs. B_w for the modulators in Table 11.3. Note that, the best trade-off among power consumption, resolution and f_n is obtained by

19 .As proposed in [25], $FOM = [\text{Power(W)} \times 10^{12}] / [2^{\text{Resolution(bit)}} \times f_d(\text{Samples/s})]$ can be used for comparative evaluation of LP $\Sigma\Delta$ Ms. This formula, which represents the energy (expressed in picojoules) needed per conversion, is not suited to compare BP $\Sigma\Delta$ Ms because it does not include f_n . However, this data constitutes one of the most important design specifications.

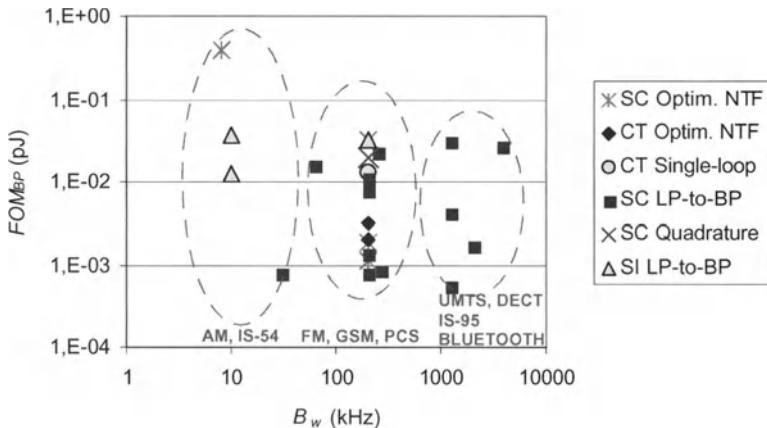


Figure 11.33: FOM_{BP} vs. B_w for the modulators in Table 12.3.

LP-to-BP BP $\Sigma\Delta$ M_s, most of them using SC circuits. However, to the best of our knowledge, only a few of them were integrated in a standard CMOS process [30][48][54]. This is due to the fact that SC circuits are very sensitive to capacitor linearity and most linear capacitor structures are formed by two poly layers – not available in standard processes.

This fact has motivated the exploration of other analog techniques compatible with standard, digital VLSI technologies. This is the case of SI circuits, which during the last few years have been used for different analog functions, including filtering and A/D conversion [18][56][57].

In the case of LP $\Sigma\Delta$ M_s, the potential advantages of SI circuits have not been demonstrated in practice. To illustrate this, Fig. 11.34 compares the FOM of both SI and SC LP $\Sigma\Delta$ M_s reported in literature. Note that the performance obtained by SI circuits is worse than that obtained by SC circuits. Observe that, while most SC modulators feature an FOM below about 10pJ, practically all SI modulators obtain an FOM above 100pJ.

However, as demonstrated in [18], the performance of SI $\Sigma\Delta$ M_s can be enhanced by means of a design methodology supported by a systematic analysis and modeling of error mechanisms. This is illustrated in Fig. 11.33, where the performance of the SI BP $\Sigma\Delta$ M_s reported in [18] – intended for digital radio receivers – is compared to that of the BP $\Sigma\Delta$ M_s in Table 12.3. It can be seen that Fig. 11.33 shows a better performance comparison between SI and SC than in LP $\Sigma\Delta$ M_s (Fig. 11.34), thus demonstrating that the SI technique is a viable alternative to the traditional SC technique for implementing BP $\Sigma\Delta$ M_s in deep submicron CMOS

technologies.

11.7 REFERENCES

- [1] A.A. Abidi: "Low-Power Radio-Frequency ICs for Portable Communications", *Proc. of the IEEE*, Vol.83, pp. 544-569, April 1995.
- [2] J. Mitola: "Technical Challenges in the Globalization of Software Radio", *IEEE Comm. Mag.*, pp. 84-89, February 1999.
- [3] B. Bing and N. Jayant: "A Cellphone for all Standards", *IEEE Spectrum*, pp. 34-39, May 2002.
- [4] H. Inose, Y. Yasuda and J. Murakami: "A Telemetering System by Code Modulation – Δ - Σ Modulation", *IRE Transactions on Space Electronics and Telemetry*, Vol. 8, pp. 204-209, September 1962.
- [5] R. Schreier and M. Snelgrove: "Bandpass Sigma-Delta Modulation", *Electronics Letters*, Vol. 25, pp. 1560-1561, November 1989.
- [6] M. Ismail: "Introduction to RF CMOS IC Design for Wireless Applications", *Tutorial 2 in the 2001 Conference on Design of Circuits and Integrated Systems*, November 2001.
- [7] A. R. Feldman, *High-Speed, Low-Power Sigma-Delta Modulators for RF Baseband Channel Applications*. Ph. D. Dissertation, University of California, Berkeley, 1998.
- [8] B. Razavi, *RF Microelectronics*. Prentice-Hall, 1998.
- [9] T.H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge University Press, 1998.
- [10] L. Breems and J.H. Huijsing, *Continuous-Time Sigma-Delta Modulation for A/D Conversion in Radio Receivers*. Kluwer Academic Publishers, 2001.
- [11] T. Burger and Q. Huang: "A 13.5-mW 185-Msample/s Δ Σ Modulator for

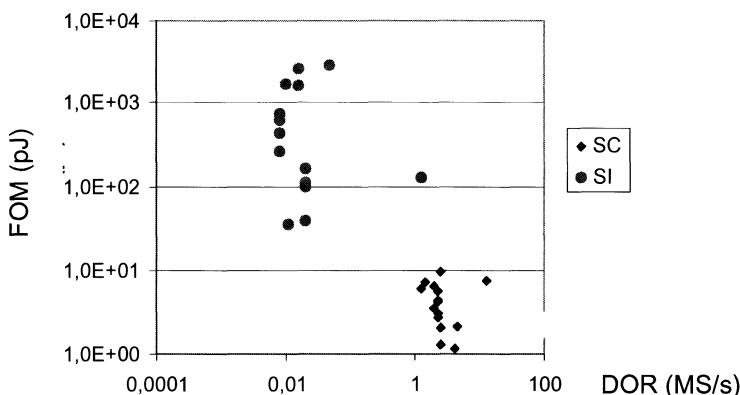


Figure 11.34: FOM vs. f_d (DOR) for both SI and SC LP Σ ΔMs.

- UMTS/GSM Dual-Standard IF Reception”, *IEEE Journal of Solid-State Circuits*, Vol. 36, pp. 1868-1878, December 2001.
- [12] S.R. Norsworthy, R. Schreier, G.C. Temes, *Delta-Sigma Converters. Theory, Design and Simulation*. IEEE Press, 1997.
 - [13] S.A. Jantzi, M. Snelgrove and P.F. Ferguson: “A fourth-order Bandpass sigma-delta Modulator”, *IEEE J. Solid-State Circuits*, Vol. 28, pp. 282-291, March 1993.
 - [14] F.W. Singor and W. M. Snelgrove: “Switched-Capacitor Bandpass Delta-Sigma A/D Modulation at 10.7 MHz”, *IEEE J. Solid-State Circuits*, pp. 184-192, March 1995.
 - [15] A.K. Ong, *Bandpass Analog-to-Digital Conversion For Wireless Applications*. Ph. D. Dissertation, Stanford University, 1998.
 - [16] R.G. Vaughan, N.L. Scott and D. Rod White: “The Theory of Bandpass Sampling”, *IEEE Trans. on Signal Processing*, Vol. 39, pp 1973-1984, September 1991.
 - [17] P.H. Gailus, W. J. Turney, and F. R. Yester: “Method and Arrangement for a Sigma Delta Converter for Bandpass Signals”, U.S. Patent 4,857,828, filed Jan. 28 1988, issued Aug. 15 1989.
 - [18] J.M. de la Rosa, B. Pérez-Verdú and A. Rodríguez-Vázquez, *Systematic Design of CMOS Switched-Current Bandpass Sigma-Delta Modulators for Digital Communication Chips*. Kluwer Academic Publishers, 2002.
 - [19] J. Van Engelen and R. Van de Plassche, *Bandpass Sigma Delta Modulators*. Kluwer Academic Publishers, 1999.
 - [20] J. A. Cherry and W. M. Snelgrove, *Continuous-Time Delta-Sigma Modulators for High-Speed A/D Conversion: Theory, Practice and Fundamental Performance Limits*. Kluwer Academic Publishers, 2000.
 - [21] R. F. Cormier, T. L. Sculley, and R. H. Bamberger: “A Fourth Order Bandpass Delta-Sigma Modulator with Digitally Programmable Passband Frequency”, *Int. Journal of Analog Integrated Circuits and Signal Processing*, pp. 217-229, 1997.
 - [22] O. Shoaei and W. M. Snelgrove: “Design and Implementation of a Tunable 40MHz-70MHz Gm-C Bandpass $\Delta\Sigma$ Modulator”, *IEEE Trans. on Circuits and Systems-II*, Vol. 44, pp. 521-530, July 1997.
 - [23] L. Louis, J. Abcarius, and G. W. Roberts: “An Eighth-Order Bandpass $\Delta\Sigma$ Modulator for A/D Conversion in Digital Radio”, *IEEE Journal of Solid-State Circuits*, pp. 423-431, April 1999.
 - [24] R. Schreier and M. Snelgrove: “Decimation for Bandpass Sigma-Delta Analog-To-Digital Conversion”, *Proc. 1990 IEEE Int. Symp. Circuits and Systems*, pp. 1801-1804, May 1990.
 - [25] F. Medeiro, B. Pérez-Verdú and A. Rodríguez-Vázquez, *Top-down Design of High-Performance Sigma-Delta Modulators*. Kluwer Academic Publishers,

- 1999.
- [26] J.C. Candy and O.J. Benjamin: "The Structure of Quantization Noise from Sigma-Delta Modulation", *IEEE Trans. on Circuits and Systems - I*, pp. 1316-1323, September 1981.
 - [27] Z.X. Zhang, G.C. Temes and Z. Czamul: "Bandpass Sigma-Delta convertor using two-path multibit structure", *Electronics Letters*, Vol.27, pp. 2008-2009, October 1991.
 - [28] S.A. Jantzi, K. W. Martin, and A. S. Sedra: "Quadrature Bandpass $\Delta\Sigma$ Modulation for Digital Radio", *IEEE J. Solid-State Circuits*, Vol. 32, pp. 1935-1949, December 1997.
 - [29] R. Gregorian, G. C. Temes, *Analog MOS Integrated Circuits For Signal Processing*. John Wiley&Sons, 1986.
 - [30] A.K. Ong and B.A. Wooley: "A Two-Path Bandpass $\Sigma\Delta$ Modulator for Digital IF Extraction at 20MHz", *IEEE Journal of Solid-State Circuit*, Vol. 32, pp. 1920-1933, December 1997.
 - [31] H. Tao, J.M. Khoury: "Analysis of Timing Jitter in Bandpass Sigma-Delta Modulators", *IEEE Trans. on Circuits and Systems-II*, pp. 991-1001, August 1999.
 - [32] O. Shoaei and W. M. Snelgrove: "A multi-feedback design for LC bandpass Delta-Sigma Modulators", *Proc. 1995 IEEE Int. Symp. Circuits and Systems*, pp. 171-173, May 1995.
 - [33] A. Hairapetian: "An 81-MHz IF Receiver in CMOS", *IEEE Journal of Solid-State Circuits*, pp. 1981-1986, December 1996.
 - [34] L. Longo and H. Bor-Rong: "A 15b 30kHz Bandpass Sigma-Delta Modulator", *Proceeding 1993 IEEE Int. Solid-State Circuit Conference*, pp. 226-227, March 1993.
 - [35] R. Schreier: "On the use of chaos to reduce idle-channel tones in delta-sigma modulators", *IEEE Trans. on Circuits and Systems - Part I*, pp. 539-547, August 1994.
 - [36] T. Salo, S. Lindfors and K. Halonen: "A 80-MHz Bandpass $\Sigma\Delta$ Modulator for a 100-MHz IF Receiver", *IEEE J. Solid State Circuits*, Vol. 37, pp. 798-808, July 2002.
 - [37] G. Raghavan, J.F. Jensen, J. Laskowski, M. Kardos, M. G. Case, M. Sokolich and S. Thomas: "Architecture, Design, and Test of Continuous-Time Tunable Intermediate-Frequency Bandpass Delta-Sigma Modulators", *IEEE J. Solid-State Circuits*, pp. 5-13, January 2001.
 - [38] W. Yu, S. Sen and B. H. Leung: "Distortion Analysis of MOS Track-and-Hold Sampling Mixers Using Time-Varying Volterra Series", *IEEE Trans. Circuits and Systems - II*, Vol. 46, pp. 101-113, February 1999.
 - [39] S.A. Jantzi, W.Martin Snelgrove, and P. F. Ferguson: "A 4th-Order BandPass Sigma-Delta Modulator", *Proc. of 1992 IEEE Custom Integrated Circuit Con-*

- ference, pp. 16.5.1-4, May 1992.
- [40] O.Norman: "A BandPass Delta-Sigma Modulator For Ultrasound Imaging at 160MHz Clock Rate", *IEEE Journal of Solid-State Circuits*, pp. 2036-2041, December 1996.
 - [41] B.S. Song: "A Fourth-Order Bandpass Delta-Sigma Modulator with Reduced Number of Op Amps", *IEEE Journal of Solid-State Circuits*, Vol. 25, pp. 1309-1315, December 1995.
 - [42] A. Tabatabaei, B. Wooley: "A Wideband Bandpass Sigma-Delta Modulator for Wireless Applications", *IEEE Int. Symp. on VLSI Circuits*, pp. 91-92, 1999.
 - [43] E. André, D. Morche, F. Balestro and P. Senn: "A 2-path $\Sigma\Delta$ Modulator for Bandpass Application", *Proc. of IEEE-CAS Workshop on Analog and Mixed IC Design*, pp. 87-91, September 1996.
 - [44] H. Liu, X. Yu, T. Sculley and R. H. Bamberger: "A Fourth Order Bandpass Delta-Sigma A/D Converter with Input Modulation Network and Digitally Programmable Passband", *Proc. 1997 IEEE Int. Symp. Circuits and System*, pp. 385-388, May 1997.
 - [45] E. André, G. Martel, D. Morche, and P. Senn, "A Bandpass A/D Converter and its I/Q Demodulator for Integrated Receivers", *Proceedings of the 24th European Solid-State Circuits Conference, ESSCIRC'98*, pp. 416-419, September 1998.
 - [46] S. Bazarjani, W.M. Snelgrove: "A 160MHz Fourth-Order Double-Sampled SC Bandpass Sigma-Delta Modulator", *IEEE Trans. Circuits and Systems-II*, Vol. 45, pp. 547-555, May 1998.
 - [47] S. Chuang, H. Liu, X. Yu, T.L. Sculley, R. H. Bamberger: "Design and Implementation of Bandpass Delta-Sigma Modulators Using Half-Delay Integrators", *IEEE Trans. Circuits and Systems-II*, Vol. 45, pp. 535-546, May 1998.
 - [48] J. Park, E. Joe, M. Choe, B. Song: "A 5-MHz IF Digital FM Demodulator", *IEEE Journal of Solid-State Circuit*, Vol. 34, pp. 3-11, January 1999.
 - [49] D. Tonietto, P. Cusinato, F. Stefani, A. Baschirotto: "A 3.3V CMOS 10.7MHz 6th-order bandpass $\Sigma\Delta$ Modulator with 78dB Dynamic Range", *Proceedings of the 25th European Solid-State Circuits Conference, ESSCIRC'99*, pp. 78-81, September 1999.
 - [50] S. Bazarjani, S. Younis, J. Goldblatt, D. Butterfield, G. McAllister, S. Ciccarelli: "An 85MHz IF Bandpass Sigma-Delta Modulator for CDMA Receivers", *Proceedings of the 25th European Solid-State Circuits Conference, ESSCIRC'99*, pp. 266-269, September 1999.
 - [51] H. Tao, J.M. Khouri: "A 400-MS/s Frequency Translating Bandpass Sigma-Delta Modulator", *IEEE J. Solid State Circuits*, Vol. 34, pp. 1741-1752, December 1999.
 - [52] A. Tabatabaei and B. Wooley: "A Two-path Bandpass Sigma-Delta Modulator with Extended Noise Shaping", *IEEE J. Solid State Circuits*, Vol. 35, pp.

- 1799-1809, December 2000.
- [53] P. Cusitano, F. Stefani and A. Baschirotto, "A 73dB SFDR 10.7MHz 3.3V CMOS Bandpass $\Sigma\Delta$ Modulator sampled at 37.05MHz", *Proc. of the 26th European Solid-State Circuits Conference*, pp. 80-83, September 2000.
 - [54] V. S. L. Cheung, H.C. Luong and W.H. Ki, "A 1V 10.7MHz Switched-Opamp Bandpass $\Sigma\Delta$ Modulator using Double-Sampling Finite-Gain Compensation Technique", *2001 IEEE Int. Solid-State Circuit Conference*, pp. 52-53, February 2001.
 - [55] T. Ueno, A. Yasuda, T. Yamaji and T. Itakura, "A Fourth-Order Bandpass $\Delta-\Sigma$ Modulator Using Second-Order Bandpass Noise-Shaping Dynamic Element Matching", *IEEE J. Solid State Circuits*, Vol. 37, pp. 809-816, July 2002.
 - [56] N.Tan, *Switched-Current Design and Implementation of Oversampling A/D Converters*. Kluwer Academic Publishers 1997.
 - [57] B.E. Jonsson, *Switched-Current Signal Processing and A/D Conversion Circuits - Design and Implementation*. Kluwer Academic Publishers 2000.

Chapter 12

Decimation Filter Design for Sigma-Delta Converters

Valentino Liberali, Roberto Rossi, and Guido Torelli

12.1 INTRODUCTION

In the field of digital signal processing, the concept of sampling a continuous-time process is fundamental. Sampling means extracting a set of numbers that in some sense represent the process that is being sampled.

If we denote a continuous-time function as $x_C(t)$, we can define the set of samples as $x_D(n)$, where the correspondence between t and n depends on the type of sampling process and can be written as $t = q(n)$. Among the many types of sampling that have been discussed in the literature, we can mention nonuniform sampling and uniform sampling [1][2].

Uniform periodic sampling is the most common form of sampling and is described by $q(n) = nT$, that is $t = nT$. The samples $x_D(n)$ are uniformly spaced by an amount equal to the sampling period T .

Obviously, $x_C(t)$ can be sampled with any sampling period. However, if unambiguous reconstruction of the continuous-time function from its sampled version $x_D(n)$ is required, T must be chosen in order to satisfy the Nyquist sampling theorem, which states that the sampling rate $F = 1/T$ must be at least twice as high as the highest frequency component of the signal $x_C(t)$.

Sometimes it is necessary to convert the sampling rate of a signal to a lower rate (*decimation*) or to a higher one (*interpolation*) [3]-[5]. The following sections will focus on decimation.

The concept of a unique waveform corresponding to a digital sequence can be useful to gain a deeper understanding of some processing algorithms. However, this correspondence with the analog world is not strictly necessary. In fact, decimation can be viewed from two different perspectives. On the one hand, decimation is a sampling rate conversion. By decimation, a continuous-time signal is resampled at a lower rate. From this point of view, the spectrum of the

signal gives a very deep insight into the process. This approach is more general, although it requires thinking in terms of a continuous-time signal.

On the other hand, decimation can also be seen as a fully discrete process. By decimation, a sequence of numbers is derived from another one. In this case, there is no analog or continuous-time counterpart. The concept of z-transform is still valid and useful, yet less insightful.

Clearly, whenever decimation is related to a continuous-time signal, the former point of view is much more convenient and should be followed. For example, decimation of a sigma-delta modulated signal requires a detailed knowledge of signal and noise spectra in order to carry out a proper design. In this case, the first approach is not only recommended but also mandatory.

Both decimation and interpolation can be formulated in terms of linear filtering operations [6][7], although, in principle, non-linear processing is also possible.

12.2 BASIC CONCEPTS OF DECIMATION

12.2.1 Decimation in the z domain

A general description of a sampling rate conversion system is given in Fig. 12.1. From the input signal x_n , sampled at the rate $F = 1/T$, we wish to compute the signal y_m with a new sampling rate $F' = 1/T'$. Although it is not strictly necessary, we will assume that the ratio of the two sampling periods can be expressed as a rational number, that is $T'/T = F/F' = M/L$.

The sampling rate converter in Fig. 12.1 is a linear time-varying system. The system transfer function $g_m(n)$ gives the response of the system at the output sample time m to an input at the input sample time $\lfloor mM/L \rfloor - n$.

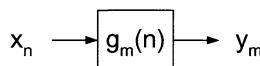


Figure 12.1: Digital sampling rate conversion.

A general expression that relates the input and the output signal is

$$y_m = \sum_n g_m(n) x\left(\left\lfloor \frac{mM}{L} \right\rfloor - n\right) \quad (12.1)$$

Decimation by an integer factor M is a special case of sampling rate conversion, where $L=1$. A signal is said to be decimated by M when one every M samples is taken. The symbol of a decimator is shown in Fig. 12.2. In this figure, the sequence x_n is decimated by M . More precisely, this operation

generates a new sequence, named y_n , which can be described as $y_n = x_{nM}$.

Usually, a filtering function is also required because aliasing must be prevented. This filtering block is what in the previous figure was named $g_m(n)$. It is not included in the symbol of Fig. 12.2, as the issue of aliasing will be addressed later.

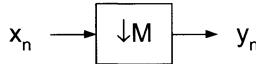


Figure 12.2: Symbol of a decimator.

It is possible to develop a useful formula to describe decimation in the z domain. Such a formula gives a deep insight on what happens when a digital sequence is decimated.

Since decimation is equivalent to resampling an analog signal, it is reasonable to expect some deep similarities between the effect of sampling a continuous-time signal and the effect that decimation has on a discrete sequence.

The z -transform of the output sequence in terms of the z -transform of the input sequence is

$$Y(z) = W(z^{1/M}) = \frac{1}{M} \sum_{p=0}^{M-1} X\left(z^{1/M} e^{-j2\pi p/M}\right) \quad (12.2)$$

It is quite evident that decimation is not a time-invariant operation.

To gain a deeper insight, we can apply a space mapping operation from the z domain to the digital frequency domain [8][9]. By setting $z = e^{j\Omega}$, the z space is mapped to the Ω space, where Ω is the normalized digital frequency ($\Omega = \omega T$). We obtain

$$Y(e^{j\Omega}) = \frac{1}{M} \sum_{p=0}^{M-1} X\left(e^{j\frac{\Omega-2\pi p}{M}}\right) \quad (12.3)$$

The above relationship can be more conveniently written in a shorter notation, where the similarities between analog sampling and digital decimation are immediately apparent:

$$Y(\Omega) = \frac{1}{M} \sum_{p=0}^{M-1} X\left(\frac{\Omega - 2\pi p}{M}\right) \quad (12.4)$$

Thus, in the frequency domain, a decimator has two concurrent effects, which are spectral replication and frequency scaling by a factor of M . Let us assume x_n is a signal whose spectral components are negligible for frequencies higher than

π/M . Fig. 12.3 shows the spectrum of such a signal in a typical case. As it always happens, in the digital domain the spectrum is periodic with a period equal to 2π . After spectral replication, the spectrum looks like in Fig. 12.4. Finally, Fig. 12.5 shows the final result of decimation after the frequency axis has been scaled.

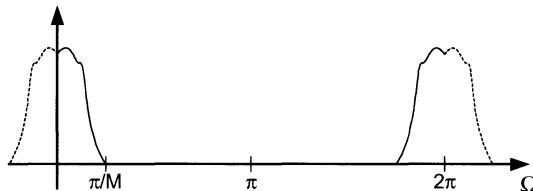


Figure 12.3: Typical spectrum of a band-limited digital signal.

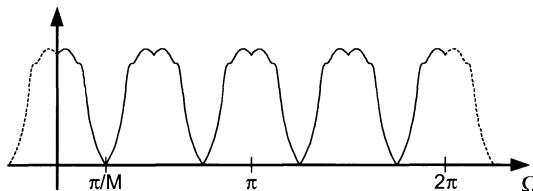


Figure 12.4: Spectrum of the signal in Fig. 12.3 after spectral replication.

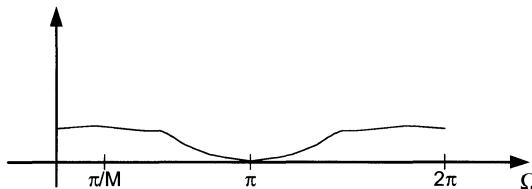


Figure 12.5: Spectrum of the signal in Fig. 12.4 after scaling the frequency axis.

To gain a deeper understanding, let us now consider the case of a signal whose spectral components extend over π/M . An example of such a signal is represented in Fig. 12.6. In this case, decimation by M causes the spectral lobes to overlap, thus giving rise to the aliasing phenomenon (Fig. 12.7).

It is clear from Fig. 12.7 that aliasing can be regarded as a source of noise. Hence, an antialiasing filter should be used in order to improve the signal-to-noise ratio (SNR). This is similar to what happens when sampling continuous-time signals. By removing all the spectral components that lie outside the interval $0 \div \pi/M$, aliasing ceases to occur. In practice, some frequency components

outside that interval will only be attenuated, thus resulting in a lower-than-ideal SNR.

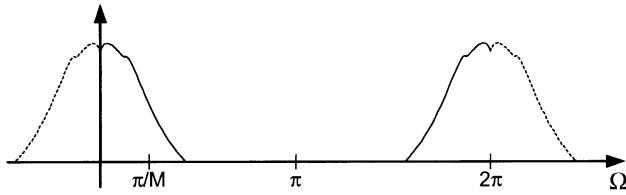


Figure 12.6: Signal whose spectrum is not limited to π/M .

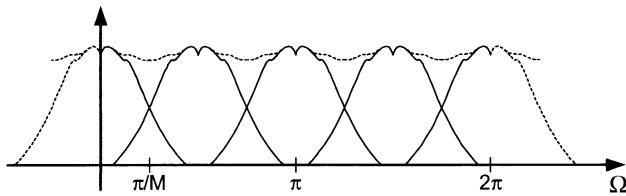


Figure 12.7: Spectrum of the signal in Fig. 12.6 after spectral replication; aliasing occurs because the spectral lobes overlap.

When sigma-delta modulation is considered, a high portion of the total amount of noise is concentrated out of the signal band. In this case, decimation causes the out-of-band spectral components of the noise to be folded onto the signal, again lowering the SNR. The antialiasing filter used for this situation is usually called a *decimation filter* and can take advantage of several optimization techniques. These issues will be addressed shortly.

Suppose we have a signal we need to decimate by M . Suppose also that the actual signal lies in the band $0 \div \pi/B$, while there is only noise from π/B to π . Only the case $B \geq M$ is of practical interest.

From the above discussion, it is clear that a low-pass antialiasing filter is needed to preserve the signal-to-noise ratio. We will now show that some intervals of frequencies do not contribute to SNR degradation because they do not get folded into the $0 \div \pi/B$ band. For this reason, in some cases the antialiasing specifications can be relaxed a lot, which can greatly simplify the design.

For a better understanding, let us first examine the case of decimation by 4 when $B = 16$ (Fig. 12.8). We will then consider the most general case.

Substituting $M = 4$, eq. (4) can be rewritten as follows.

$$Y(\Omega) = \frac{1}{4} \sum_{p=0}^3 X\left(\frac{\Omega}{4} - \frac{\pi}{2} p\right) \quad (12.5)$$

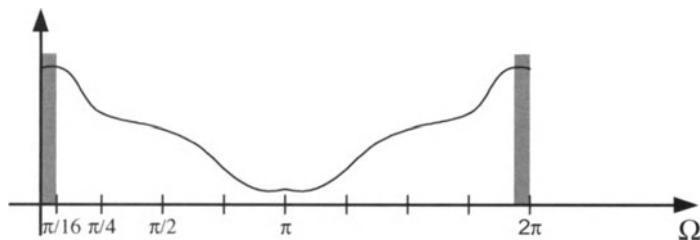


Figure 12.8: A generic signal plus noise. The shaded areas represent the signal bandwidth.

Through decimation by 4, the signal band will be stretched from 0 to $\pi/4$. To find out the spectral components of X that will contribute to the decimated signal Y , we have to look at the above formula closely. Recalling that $Y(\Omega)$ is periodic with a period of 2π , the signal after decimation must be in the intervals $-\pi/4 + 2k\pi \leq \Omega \leq \pi/4 + 2k\pi$ (where k is integer). The question which parts of the spectrum of X are taken can be easily answered if we define $\Omega' = \frac{\Omega}{4} - \frac{\pi}{2} p$, where p is an integer. We obtain

$$-\frac{\pi}{16} + \frac{\pi}{2}(k-p) \leq \Omega' \leq \frac{\pi}{16} + \frac{\pi}{2}(k-p) \quad (12.6)$$

Therefore, all the intervals that are centered at every integer multiple of $\pi/2$ and are $\pi/8$ wide affect the signal band of the decimated signal Y . A picture showing all these intervals is given in Fig. 12.9.

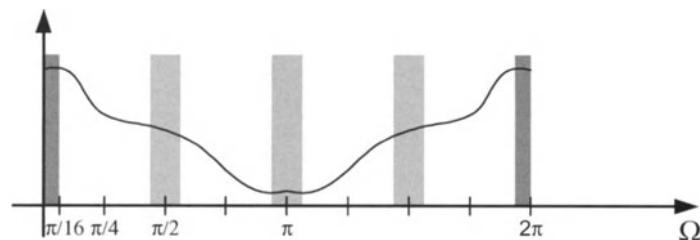


Figure 12.9: The light shaded areas mark the frequency intervals that get aliased onto the darker areas (the signal).

All the white areas are don't-care bands, which means they are not relevant from the point of view of aliasing. Hence, when designing an antialiasing filter for decimation by 4 with $B = 16$, only the noise lying in the gray intervals should

be filtered away. The presence of don't-care bands in a filter mask can simplify the design very much. This is especially useful in the first stages of a decimator, where the data rate is still high and B is considerably greater than M .

The most general case can be dealt with in the same way. From eq. (12.4) we can define the variable Ω' as $\Omega' = (\Omega - 2\pi p)/M$. Then, remembering that decimation stretches the signal bandwidth by M from π/B to $\pi M/B$, starting from the relationship $-(\pi M/B) + 2k\pi \leq \Omega \leq (\pi M/B) + 2k\pi$, we obtain

$$-\frac{\pi}{B} + \frac{2\pi}{M}(k-p) \leq \Omega' \leq \frac{\pi}{B} + \frac{2\pi}{M}(k-p) \quad (12.7)$$

The above formula can be expressed in a shorter notation as follows:

$$-\frac{\pi}{B} + \frac{2\pi k}{M} \leq \Omega' \leq \frac{\pi}{B} + \frac{2\pi k}{M} \quad (12.8)$$

Clearly, when B and M are equal, there are no don't-care bands. All frequency components get aliased onto the signal. This is what usually happens in the last stage of a multi-stage decimator, as we will see later.

Instead, when $B > M$, which happens in all the other stages of a decimator, there is some space left for optimization, as $\lceil M/2 \rceil$ don't-care bands appear in the range $0 \div \pi$.

12.2.2 The commutative rule

When dealing with decimating blocks, it is convenient to keep some useful relationships at hand. One of these is the commutative rule, first proposed by Chu and Burrus [10]. This rule explains how to swap decimating and filtering blocks.

If a sequence of samples x_n has to be decimated by a factor of M and, subsequently, filtered by $H(z)$, the commutative rule states that the same output sequence can be obtained by first filtering by $H(z^M)$ and, subsequently, by decimating by the same factor M . This property can be easily proven by analyzing the two configurations in Fig. 12.10.

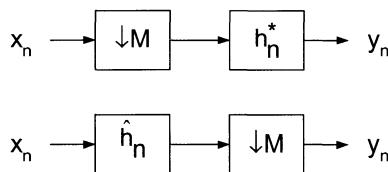


Figure 12.10: The commutative rule.

Let us assume that the following relationships hold between the two impulse responses shown in Fig. 12.10.

$$\begin{cases} \hat{h}_n = h_n \sum_k \delta(n - kM) \\ h_n^* = \hat{h}_{nM} = h_{nM} \end{cases} \quad (12.9)$$

Eqs. (12.9) state that h_n^* is a decimated version of \hat{h}_n , while \hat{h}_n is an interpolated version of h_n^* . In other words, \hat{h}_n can be obtained by adding $M - 1$ null samples between any two consecutive samples of h_n^* .

It is quite easy to show that, given the same input sequence x_n , the two systems in Fig. 12.10 yield the same output sequence y_n . In fact, the output from the first system can be expressed as follows:

$$y_n = h_n^* * x_{nM} = \sum_k h_k^* x_{nM-kM} = \sum_k h_{kM} x_{nM-kM} \quad (12.10)$$

On the other hand, the output from the second system is:

$$\begin{aligned} y_n &= (\hat{h}_n * x_n)|_{nM} = \sum_k \hat{h}_{n-k} x_k|_{nM} = \sum_k \hat{h}_{nM-k} x_k = \\ &= \sum_k x_k \sum_r h_{nM-k} \delta(nM - k - rM) = \sum_r h_{rM} x_{nM-rM} \end{aligned} \quad (12.11)$$

This proves the two systems are perfectly equivalent.

The commutative rule can also be stated in the z domain as $\hat{H}(z) = H^*(z^M)$. In fact, remembering that $\hat{h}_n = 0$ if $n \neq kM$, we may write:

$$H^*(z^M) = \sum_n h_n^* z^{-nM} = \sum_n \hat{h}_{nM} z^{-nM} = \sum_n \hat{h}_n z^{-n} = \hat{H}(z) \quad (12.12)$$

12.3 DESIGN FLOW

The design of a digital system usually consists of a number of different steps. A decimation filter represents no exception to this rule. In this paragraph, the design flow adopted for such a filter will be outlined.

First of all, a set of specifications should be taken. Next, the system should be partitioned and a suitable architecture should be determined. The coefficients of the filtering stages must be chosen accordingly. This step may require a number of iterations to achieve a satisfactory trade-off between complexity and performance.

Once the architecture and the coefficients have been defined, the obtained system must be translated into a hardware description language (HDL) in order

to proceed with the design flow. A number of HDL languages are available, among which the most popular are Verilog and VHDL.

After the HDL code has been simulated and verified, synthesis can be carried out. By synthesis, the HDL description is mapped to a gate-level description given a set of timing and design constraints. A gate-level description, also called *netlist*, specifies how the available library cells, mostly logic gates and sequential elements, must be connected to each other in order to achieve the desired circuit behavior. Timing constraints put limits on path delays, while design constraints set an upper bound on area or power consumption or, for example, introduce some requirements on the driving capability of the output ports.

In this phase, static timing analysis (STA) is used to verify whether or not the specifications are met. A few iterations may be necessary. The result of synthesis is a VHDL or Verilog netlist that can be simulated.

If STA and simulation results are correct, the netlist can go through the place-and-route (P&R) phase, whose result is the final layout. This phase may require a number of iterations, especially when the place-and-route tools are not timing driven. In this case, back-annotated verifications must be carried out in order to determine if the layout meets the target specifications.

Newer tools are timing driven instead, that is they estimate interconnection delays while doing placement and then evaluate them during the routing phase. The tool itself manages any necessary iteration automatically. The effect is a much shorter design cycle as well as a better layout.

Once the layout is finished, like for any integrated circuit, the fabrication process is launched.

12.4 ARCHITECTURE

A number of FIR filter architectures have been developed and are described in the literature [11]-[13]. Among these, we can mention the *direct* and the *transposed* forms (also called *canonical*), as well as the *systolic* and the *hybrid* forms.

Fig. 12.11 and Fig. 12.12 show the direct and the transposed form, respectively. In high-speed applications, both of them suffer from important drawbacks. The critical path of the direct form corresponds to the time required to carry out one multiplication and all the associated add operations. The ensuing delay is an increasing function of the number of coefficients and must be less than a clock period in order not to cause timing violations. The transposed form solves this problem by inserting intermediate delays between any two cascaded adders. In this case, the critical path only includes one multiplication and one addition. However, for filters with a high number of taps, the input bus capacitance can become too high, thus limiting overall performance.

To make the computation delay independent of the number of coefficients, systolic architectures can also be used. These are modular architectures where

additional pipeline delays are included in order to minimize critical paths.

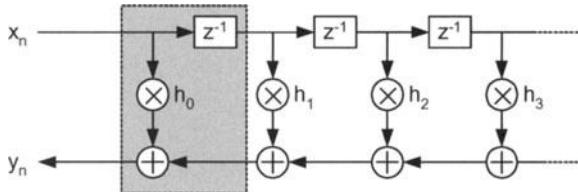


Figure 12.11: Direct form of an FIR filter. This is a canonical implementation.

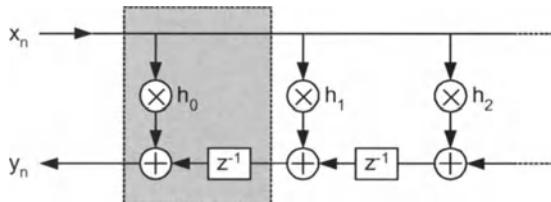


Figure 12.12: Transposed form of an FIR filter. This is canonical also.

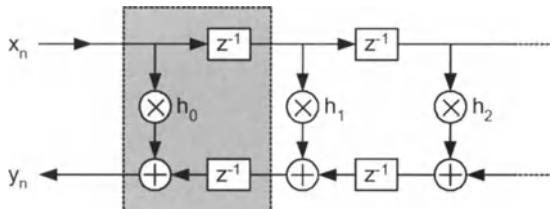


Figure 12.13: Systolic implementation of an FIR filter. The clock frequency must be doubled with respect to the case of a canonical implementation in order to keep the same throughput.

An example is given in Fig. 12.13. Like in the transposed form, the critical path only flows through one multiplication and one addition, but now the input bus capacitance is kept low. However, there is one flaw that may impact high-speed operation negatively. The z -transform of the output can be written as follows:

$$Y(z) = X(z) \sum_k h_k z^{-2k} \quad (12.13)$$

From eq. (12.13) it is seen that the clock frequency must be doubled with respect to the case of a canonical implementation if the same throughput is desired. This can be a serious problem when the clock rate is already high,

because the constraints on critical paths are tightened also. For this reason, this architecture is less attractive.

It is also possible to derive a pipelined structure that does not display this problem. A hybrid form is a mixture of the direct and the transposed form. An example of a hybrid form (*hybrid form I*) is given in Fig. 12.14.

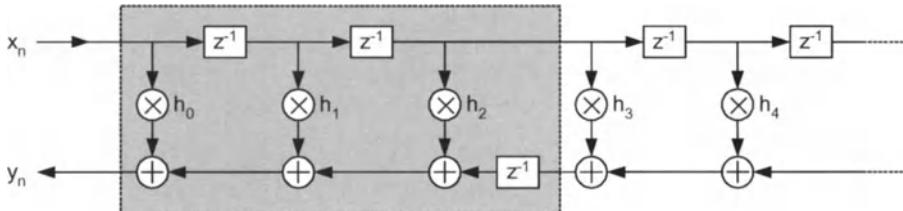


Figure 12.14: An example of hybrid form I.

Hybrid forms are modular in that they consist of the cascaded identical stages, like the other architectures shown above. They have zero latency and also have fewer registers than the systolic structures. Hybrid forms are a good compromise between the direct and transposed forms, and are therefore a very good choice for high-speed low-power applications.

12.4.1 Multi-stage vs. single-stage

When designing the architecture of a decimation filter, the way the system is partitioned is one of the most important issues. As we will show shortly, the number of stages in which the filter is divided affects area and power consumption heavily. Hence, before dealing with the architecture of each filtering stage, the overall architecture of the decimator must be determined.

Recalling the commutative rule, it is clear that, given any multi-stage decimator, it is possible to convert it into a single-stage one. This proves that a decimation filter may be partitioned in more than one way, which poses the problem of choosing the ideal partitioning strategy.

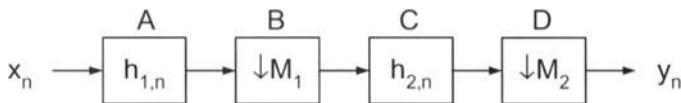


Figure 12.15: A two-stage decimator.

As an example, let us see how a two-stage design (Fig. 12.15) can be converted into a different scheme. By applying the commutative rule, blocks B and C can be exchanged, resulting in the block diagram of Fig. 12.16.

Clearly, $\hat{h}_{2,n}$ is obtained by inserting $M_1 - 1$ null samples between any two

samples of $h_{2,n}$, as stated by the commutative rule. In other words, one gets $\hat{h}_{2,n}$ by interpolating $h_{2,n}$ by a factor of M_1 . Then, discrete convolution can be used to combine the two filtering blocks together. Next, the decimation switches can be replaced with a single one whose decimation ratio is the product M_1M_2 . The result is shown in Fig. 12.17.

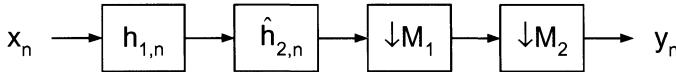


Figure 12.16: The two-stage decimator of Fig. 12.15 after application of the commutative rule.

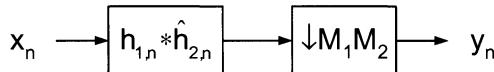


Figure 12.17: Single-stage decimator equivalent to the one depicted in Fig. 12.16.

More in general, if the number of stages is greater than two, by repeatedly applying the commutative rule all the filtering blocks can be grouped together at the left end of the data path, while all the decimation switches are moved toward the right end. Then, the group of filtering blocks can be replaced with a single filter by discrete convolution. Similarly, a single block can be substituted for the whole chain of decimation switches by simply multiplying the decimation ratios. This simple sequence of operations allows any multi-stage design to be converted to a single-stage one.

Starting from the above considerations, we will now show that a multi-stage decimator always outperforms a single-stage design in terms of both silicon area and computational complexity, measured as the total number of coefficients and the number of operations per output sample, respectively.

If the number of coefficients of the two filters h_1 and h_2 are denoted with N_1 and N_2 , respectively, the interpolated impulse response $\hat{h}_{2,n}$ has N_2M_1 coefficients (N_2 of which are different from zero). Therefore, the length of the discrete convolution $h_{1,n} * \hat{h}_{2,n}$ is not greater than $N_1 + N_2M_1$. Thus, it can be said that, at least to a rough approximation, the required silicon area is proportional to that number, that is

$$\text{Area} \propto N_1 + N_2M_1 \quad (12.14)$$

while for the two-stage configuration

$$\text{Area} \propto N_1 + N_2 \quad (12.15)$$

As to computational complexity, a rough estimate can be quickly given if we count the total number of MAC (multiply-and-accumulate) operations required per output sample. Noticing that the number of MAC operations required to filter a sequence of samples is equal to the impulse response length, we may write

$$N_{\text{MAC, Two-stage}} = M_1 M_2 N_1 + M_2 N_2 \quad (12.16)$$

while for the equivalent single-stage configuration

$$N_{\text{MAC, Equivalent}} = M_1 M_2 (N_1 + N_2 M_1) \quad (12.17)$$

which can be considerably higher. In fact, we have:

$$N_{\text{MAC, Equivalent}} = N_{\text{MAC, Two-stage}} + M_2 N_2 (M_1^2 - 1) \quad (12.18)$$

The performance improvement of the two-stage design over its equivalent single-stage counterpart can be of great significance. For example, if we need to decimate by a factor of 4, the only option is to set $M_1 = M_2 = 2$. In the case $N_1 = 50$ and $N_2 = 100$, we have

$$\begin{cases} \text{Two - stage} & \text{Equivalent One - stage} \\ \text{Area} \propto 150 & \text{Area} \propto 250 \\ N_{\text{MAC}} = 400 & N_{\text{MAC}} = 1000 \end{cases} \quad (12.19)$$

Therefore, splitting a decimator into two stages makes it possible to save a lot of area and computational power.

The above equations can easily be extended to the most general case of an S -stage decimator. In this case, we obtain

$$\begin{cases} \text{Area} \propto \sum_{s=1}^S N_s \\ N_{\text{MAC}} = \sum_{s=1}^S \left(N_s \prod_{k=s}^S M_k \right) \end{cases} \quad (12.20)$$

while for the equivalent one-stage decimator we have:

$$\begin{cases} \text{Area} \propto \sum_{s=1}^S \left(N_s \prod_{k=1}^{s-1} M_k \right) \\ N_{\text{MAC}} = \left[\sum_{s=1}^S \left(N_s \prod_{k=1}^{s-1} M_k \right) \right] \prod_{s=1}^S M_s \end{cases} \quad (12.21)$$

For example, let us consider a three-stage decimator ($S = 3$) whose

decimation ratio of 16 can be factored as $4 \times 2 \times 2$ ($M_1 = 4, M_2 = M_3 = 2$). Let us assume that $N_1 = 20$, $N_2 = 15$ and $N_3 = 100$. Then, by substituting into eqs. (12.20) and (12.21), we have:

$$\left\{ \begin{array}{l} \text{Three - stage} \\ \text{Area} \propto 135 \\ N_{\text{MAC}} = 3200 + 60 + 200 = 3460 \end{array} \right. \quad \left\{ \begin{array}{l} \text{Equivalent One - stage} \\ \text{Area} \propto 20 + 60 + 800 = 880 \\ N_{\text{MAC}} = 16 \cdot 880 = 14080 \end{array} \right. \quad (12.22)$$

Compared to the equivalent single-stage design, the three-stage decimator allows reducing the required silicon area by 6.5 and the computational power by 4.

In general, the greater the number of stages, the higher the gain in terms of both area and computational power. However, for a high number of stages, this might in practice no longer be true because of the increasing complexity of the required control logic. In fact, a multi-stage decimator requires additional control circuitry compared to a single-stage one. There exists a point where the advantage coming from an additional decimation stage is canceled out by the increased complexity of the control blocks. Anyway, when the overall decimation ratio is not very high, it is usually better to split the design into as many stages as possible.

In the following of this section, we will focus on the particular case of a three-stage decimation filter. The stages are referred to as sinc filter, equalizer, and band filter, respectively. The reason for these names will be clarified.

12.4.2 First stage

In our example, despite the low number of coefficients (20) the first stage has a heavy impact on the total number of MAC operations (3200). As the first stage works at the highest data rate, most of the computational power is concentrated in it. For this reason, a standard FIR filter architecture is not considered adequate for the first stage of a decimator, which then usually implements an entirely different architecture. In fact, a non-flat in-band frequency response can be easily allowed for and, hence, equalized in the following stages if, as a result of this choice, the total computational power can be reduced by roughly an order of magnitude. In general, a sinc filter is chosen as the first stage because it can be conveniently implemented in a very efficient manner.

Furthermore, in the first stages the antialiasing requirements are highly relaxed because $B > M$ in eq. (12.8), which gives the designer a chance to trade frequency distortion for lower power. This is exactly the purpose of a sinc filter. A limited amount of frequency distortion is introduced into the signal band, as this enables the use of a very efficient architecture. Then, this distortion has to be equalized in the next stages of the design, but the cost of this operation is usually almost negligible.

Filters with a $\text{sinc}^K(f)$ response are appropriate for decimating a sigma-delta modulated signal down to four times the Nyquist rate. Further decimation usually requires filters that cut off more sharply at the edge of the pass-band [14].

Let us consider the discrete convolution among K identical N -tap comb filters. The impulse response of such a filter can be written as

$$h_{\text{sinc}^K}(n) = (h_{\text{comb}}(n))^{*K} \quad (12.23)$$

Correspondingly, its z -transform is

$$H_{\text{sinc}^K}(z) = \left(\frac{1}{N} \sum_{k=0}^{N-1} z^{-k} \right)^K = \left(\frac{1}{N} \frac{1-z^{-N}}{1-z^{-1}} \right)^K \quad (12.24)$$

Again, if we substitute $z = e^{j\Omega}$, we obtain the frequency response

$$H_{\text{sinc}^K}(\Omega) = \left(\frac{1}{N} \frac{1-e^{-jN\Omega}}{1-e^{-j\Omega}} \right)^K = e^{-j\frac{\Omega}{2}K(N-1)} \left(\frac{\sin \frac{N\Omega}{2}}{N \sin \frac{\Omega}{2}} \right)^K \quad (12.25)$$

In the low frequency region, the response can be approximated as

$$H_{\text{sinc}^K}(\Omega) \approx e^{-j\frac{\Omega}{2}K(N-1)} \text{sinc}^K\left(\frac{N\Omega}{2}\right) \quad (12.26)$$

which explains the name sinc^K given to this type of filters.

From the expression of the frequency response, it is seen that an N -tap sinc^K filter displays the same set of lobes and zeros as its N -tap comb filter counterpart.

12.4.2.1 Decimation with a sinc filter

Suppose we want to use a sinc^K filter in the first decimation stage for an L -order sigma-delta modulator. The sinc response in the signal band introduces a linear distortion that needs to be corrected in some way. Anyway, this response can be very easily equalized in the next stages of the decimator.

The real question is about the amount of noise that gets folded into the signal band. Fortunately, not all the noise power gets aliased onto the signal, as in the stage considered the decimation factor is only a small fraction of the whole oversampling ratio of the sigma-delta modulator. In other words, B is considerably greater than M .

The knowledge of the decimation factor of the first decimating stage (M_1) and the oversampling ratio of the modulator (B) allows defining the filter mask of

the first decimating stage as a set of don't-care bands and a set of bands where the amplitude of the response must be null, as stated in eq. (12.8). Obviously, to fit a sinc^K filter response into such a mask, the zeros of the sinc filter must lie in the null bands of the mask. The number of taps of the sinc^K filter is fixed by this constraint and must be equal to the decimation ratio of the stage.

Furthermore, the order K of the sinc^K filter is related to the antialiasing performance of the stage. Indeed, the higher K , the better the noise cancellation behavior around the filter zeros. Nevertheless, a higher value of K introduces a sharper distortion in the signal band. Anyway, this is of lesser importance compared to the aliasing constraint.

It is clear that K directly affects the maximum signal-to-noise ratio of the sigma-delta converter. In principle, we could imagine using an ideal low-pass filter in the subsequent stages of the decimator. This would propagate the same signal-to-noise ratio present at the output of this stage to the output of the last decimating stage. However, since an ideal filter is impractical, in a real situation all the other stages also contribute to lower the signal-to-noise ratio. In other words, each stage introduces some noise into the signal band, thus lowering the overall signal-to-noise ratio.

These arguments show that there is a lower bound to the value of K .

12.4.2.2 An efficient architecture for the sinc filter

One method to make the filter efficient is to design the transfer function $H(z)$ so that it has the form

$$H(z) = f(z)g(z^{M_1}) \quad (12.27)$$

where the factor $g(z^{M_1})$ can be implemented at a lower rate as $g(z)$ as shown in Fig. 12.18.

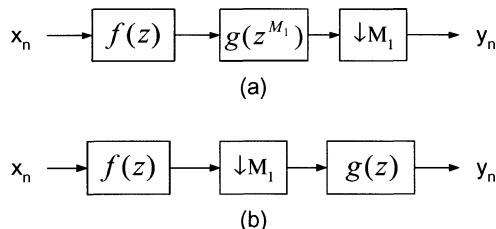


Figure 12.18: (a) A filter that can be made more efficient; (b) the same filter after applying the commutative rule.

By this method, a high-order filter $H(z)$ can be split into two lower-order filters. The arithmetic rate, the number of coefficients and the number of registers

are reduced roughly by M_1 .

One such example is a cascade of comb filters, that is a sinc^K filter, which can also be written as $\left(\frac{1-z^{-M_1}}{1-z^{-1}}\right)^K$. In this case, $g(z^{M_1}) = (1-z^{-M_1})^K$.

12.4.3 Equalizer

The attenuation introduced by the sinc filter is generally compensated in the second stage, which has a frequency response that approximates $(H_{\text{sinc}^K}(\Omega))^{-1}$ in the signal band. This stage is usually implemented with a multiplierless FIR filter.

12.4.3.1 The CSD representation

The canonical signed-digit (CSD) representation of the filter coefficients allows minimizing the number of operations required to implement the discrete convolution between the input sequence and the impulse response [15]. It can be viewed as an extension of the standard binary representation of numbers. Unlike the binary representation, which only uses zeros and ones as digits, the signed-digit (SD) representation can use 0, +1 and -1. SD is a redundant representation, as any given number can be written in a number of different ways.

For example, consider three possible representations of the number 7:

$$\begin{array}{cccccc} 0 & 0 & +1 & +1 & +1 \\ 0 & +1 & 0 & 0 & -1 \\ +1 & -1 & 0 & 0 & -1 \end{array}$$

Among all the possible representations of any given number, those with the smallest number of non-null digits are called *canonical*. The CSD representation of 7 is thus “0+00-” (i.e., 0, +1, 0, 0, -1).

Multiplication by a CSD-encoded coefficient only involves additions, subtractions and shifts. If the digit is +1, the shifted operand is added. Conversely, if the digit is -1, a subtraction is performed. Obviously, if the digit is 0, no operation is performed on the corresponding operand. We can then conclude that the number of operations is minimized and there is no need for general-purpose multipliers.

Thanks to the CSD representation of the filter coefficients, the number of operations per unit time performed by the overall decimator can be greatly reduced compared to the case when the natural binary representation is used.

12.4.3.2 Architecture

It should be clear that, in a multiplier-free implementation, multipliers are realized with add and subtract operations. Of course, shifts are also needed, but

they do not cause any waste of area, as shifting is simply a matter of wiring. Compared to the cases of Fig. 12.11, Fig. 12.12, Fig. 12.13, and Fig. 12.14, more adders are needed because *each* multiplier has to be substituted with *a number of* adders. Thus, it is convenient to arrange all these add and subtract operations in a tree, in order to reduce their number and save even more area. Unfortunately, any two arithmetic blocks may not be arranged in the same tree if there is a pipeline register between them, unless, obviously, more pipeline registers are added in order to balance the delays. Therefore, a good choice is to use the direct form of Fig. 12.11 and adapt it to the specific needs.

A possible architecture can be defined as a *multiplier-free pipelined direct form*. An example of such a structure is represented in Fig. 12.19 for a very short impulse response. Latency is not zero, as it would be with a hybrid form, but is proportional to the logarithm of the number of taps. Thus, this topology is still a good compromise between canonical and hybrid forms.

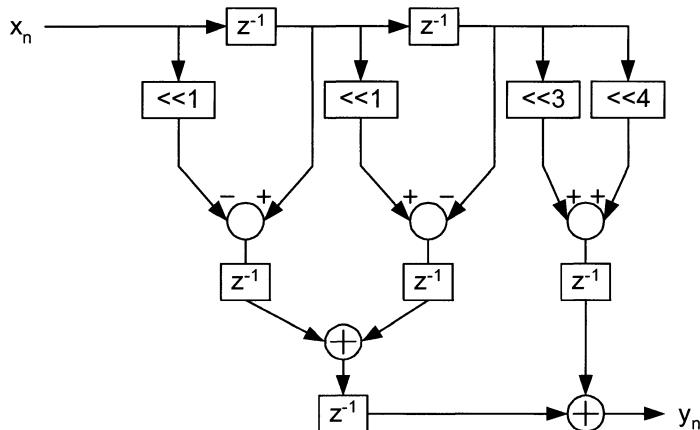


Figure 12.19: Multiplier-free pipelined direct form. In this example, the impulse response has three coefficients: -2 , 3 , and 23 . The expression “ $<< n$ ” means “left-shift by n bits”.

12.4.4 Band filter

In the last stage of the decimation filter, both the decimation ratio and the oversampling ratio have the same value ($M = B$). Recalling eq. (12.8), there are no don't-care bands, so this is a very selective filter. The passband of this stage is equal to the signal band, hence the name given to it. Like the equalizing stage, the same multiplier-free pipelined direct form FIR filter architecture can be adopted, along with the CSD representation of the coefficients.

12.5 DESIGN ENTRY

Once the basic architectures and all the coefficients have been chosen, the design entry phase can be started. This is typically carried out by means of a hardware description language (Section 12.3).

Before starting the actual design entry phase, a few additional issues must be considered: hierarchy, granularity, and clocking strategy.

12.5.1 Hierarchy

It is well known that any synthesis tool will work better and faster if the design is hierarchical and well organized. It is easier to translate smaller design units into a gate-level description than to have a single huge design mapped to a netlist. Moreover, if the hierarchy is well designed, the synthesis tools can take advantage of repeated blocks. If a block is repeated many times and the timing constraints allow it, it can be synthesized only once. This can speed up synthesis very much.

Besides, a hierarchical design is much easier to maintain and simulate.

12.5.2 Granularity

A great number of HDL descriptions are usually possible for any given design. Hierarchy may vary very much and the blocks at any level of hierarchy can be described in several different ways. The level of detail of an HDL description greatly affects the final result of the synthesis phase. A designer may prefer a detailed description, while another may prefer a coarser one. Both styles are acceptable when simulation is the only concern.

VHDL is a very powerful programming language. Synthesizable VHDL is a subset of VHDL that has to be used when designing real-world digital systems. When non-synthesizable constructs are used, synthesis is impossible. Nonetheless, even when only synthesizable VHDL constructs are adopted, synthesis is not guaranteed to succeed. It is a matter of complexity and broadness of the design space.

The effect of a more detailed HDL description is less freedom left to the synthesis tool. This way, a smaller design space needs to be explored and synthesis is more likely to succeed.

For example, a filter could be described by a few lines of VHDL code such as:

```
y <= x - 3*xOld - 3*xOld2 + xOld3
```

and the rest of the VHDL code would define the signals in such a way that the above line is equivalent to $y_n = x_n - 3x_{n-1} - 3x_{n-2} + x_{n-3}$.

Of course, the result of synthesis would be nearly unpredictable. There is no guarantee that synthesis will succeed. Moreover, the result might be unnecessarily power or area consuming. Such an HDL description would give the designer no control over what the synthesis tool is doing. The opposite approach is a very detailed HDL code, where even the internal structure of registers and adders is meticulously described. In this second case, much tighter timing, area and power constraints may be fulfilled, but the HDL code is much bulkier, as well as lengthy to write and simulate.

A designer always has to choose the granularity of the design anywhere between these two far ends. This choice affects the quality of the synthesized design.

12.5.3 Clocking strategy

The clock nets usually have a very high fan-out, as every register has a clock input. Therefore, such nets must be handled with great care. Clock skew problems may introduce timing violations in large portions of the design. For this reason, a careful design of the clock nets is mandatory.

The design of the clock tree must be carried out during the layout phase, as only in this phase all the interconnection delays can be accurately estimated.

In a decimation filter, not all the registers have to switch at every clock cycle. Two opposite clocking strategies are therefore possible.

The easiest one is to use a single clock for all the registers and then use control signals to activate them only when necessary. This way, only one clock tree is required and synthesis is easier.

The most complex technique is to generate as many clock signals as needed in order to have each register working at the lowest possible clock frequency, thereby saving power. This strategy requires having multiple clock domains, each one with its own clock tree. The skew within each clock tree can be guaranteed to be lower than a predetermined limit, but nothing can be assumed between any two clock domains. In this case, a special clock domain interfacing technique has to be used and synthesis poses some additional concerns.

12.6 SYNTHESIS

By synthesis, a behavioral HDL code is mapped to a gate-level HDL code, also called a netlist. A standard-cell library contains all the available cells for a given integration technology.

In order for synthesis to produce a meaningful result, a set of constraints has to be provided.

12.6.1 Constraints

When a design is being optimized, two types of constraints are used: design rule constraints and optimization constraints.

Design rule constraints are implicit constraints and are defined by the technology library chosen. These constraints are requirements for a design to work correctly, and apply to any design using the library. For example, fanout, transition time and capacitance are design rule constraints.

Optimization constraints are explicit constraints and are defined by the designer.

Input and output delays are used to model external delays outside the chip. They are relative to the active clock edge. The input delay is the propagation delay of the external input path. For instance, a 6-ns input delay means that the chip input signal will change 6 ns after the active clock edge. Conversely, the output delay is the propagation delay of the external output path. For instance, a 4-ns output delay implies that the chip output signal must be ready and settled at least 4 ns before the next active clock edge.

12.6.2 Synthesis strategies

Different pieces of a design may require different compilation (synthesis) strategies. The top-down strategy is the most convenient and practical one from the designer's perspective. Unfortunately, since the whole design is treated as a single unit, this approach requires a large amount of memory and, hence, is only applicable to small designs.

The converse is to use the bottom-up compilation strategy, which deals with small or precompiled blocks, each one taken as a single unit. However, this requires very detailed information about timing, as all the sub-blocks have to be constrained individually. This is clearly not practical, unless most sub-blocks are instantiated a very large number of times.

The best strategy, in terms of both used memory and compilation time, is a combination of the two above approaches. All the repeated sub-blocks are constrained and compiled individually and then the rest of the hierarchy is synthesized in a top-down manner, keeping the resulting timing behavior and load information into account. In this way, if a register is instantiated, say, 1,000 times, it will be compiled only once. The resulting gate-level description will be preserved in the subsequent top-down phase.

This approach is applicable to any type of design and is not limited by memory. It may require some iterations until the interfaces between blocks are stable. Moreover, some manual control is generally needed.

12.6.3 Gate-level simulations

To verify the result of synthesis, gate-level simulations can also be run. Actually, such kind of simulation is not particularly important, as at this point no interconnection delays are available yet, and it is well known that in deep submicron technologies interconnection delays are at least as important as cell delays.

12.6.4 Layout

The layout phase concludes the design process of the digital filter. This phase may be carried out either manually or, more usually, by means of automatic tools. Post-layout timing verifications are mandatory, especially in the case of complex designs.

12.7 A DESIGN EXAMPLE

We will now briefly examine a practical case of a decimator that has been implemented in 0.25- μm CMOS technology. The techniques described above were applied.

The decimation filter in this example was designed to be cascaded to a fourth-order multi-bit multi-stage (2-1-1) sigma-delta modulator. The nominal input rate of the filter is 64 MHz and the decimation ratio is 16. Therefore, the resulting output rate is 4 MHz. The desired signal-to-noise ratio (SNR) is 78 dB. This value affects the design of the filter coefficients, which is omitted from this section for brevity. Also, this integrated circuit contains the noise cancellation block required by the multi-stage architecture, but this block will not be described here because this chapter only focuses on decimation filters.

12.7.1 Top level

In order to reduce area and power, the decimation ratio of 16 was factored as $4 \times 2 \times 2$. Consequently, the filter was partitioned into three blocks: the sinc filter, the equalizer and the band filter.

This integrated circuit has 16 input ports and 25 output ports. Table 12.1 lists port names and directions. The purpose of signals x (input), y (output), rst (reset) and clk (clock) should be evident. The signal *NandTreeOut* is the output of a NAND tree cell, which helps in testing the input pad cells. The signal *scanEnable* is intended for scan path testing. The signals *SincX*, *EqX*, *BandX* and *SelY* enable testing of subsets of the whole filter (see below).

In order to reduce power dissipation as much as possible, a strategy similar to clock gating was adopted [16]-[18]. No enable signals are used throughout the filter, but all the registers receive a clock signal whose frequency is as low as

possible. In other words, for every register, the clock frequency matches the output rate. For example, if the output rate of a pipeline register is 8 MS/s, its clock has a frequency of 8 MHz. In this way, the waste of power in the clock distribution network is kept to a minimum.

Table 12.1: Port names and directions.

Port name	Bus width	Direction
x	8	IN
rst	1	IN
clk	1	IN
scanEnable	1	IN
SelY	2	IN
SincX	1	IN
EqX	1	IN
BandX	1	IN
y	24	OUT
NandTreeOut	1	OUT

Of course, multiple clocks are needed throughout the data path. All the necessary clocks are generated internally by means of a clock divider. Clearly, the adopted approach poses the problem of interfacing two clock domains. This is a rather delicate matter, as the risk of timing violations at the interface between two domains is very high.

When multiple clock domains are present in the system, two types of clock skew problems arise. The first one is the clock skew within one clock domain and is referred to as *intra-clock skew*. Given a clock domain, the intra-clock skew is the maximum skew between any two points of the clock distribution network.

The second type of skew is defined *inter-clock skew*. Given two clock domains, the inter-clock skew is the maximum skew between a point of the first clock network and a point of the second one. This problem only arises when the design has two or more clock domains. Unfortunately, the inter-clock skew is much more difficult to control and, hence, special countermeasures must be adopted.

The filter of this example is affected by both the intra-clock and the inter-clock skew. The way the two types of skew can be faced is utterly different. In our case, the filter needs five clocks, whose frequencies are 64 MHz, 32 MHz, 16 MHz, 8 MHz and 4 MHz, respectively. Correspondingly, the design has five clock trees. With present CAD synthesis tools, the skew inside any given tree is guaranteed to be lower than a predetermined limit. Therefore, this problem is automatically handled during synthesis. It is treated like any other timing constraint, except that it is set on a clock port and, hence, affects all the sequential elements in the design.

As to the inter-clock skew, the situation is entirely different, as no information is available about the behavior of any given clock tree with respect to all the others. The layout tools used for our design were not able to control the skew between two clock domains. As a matter of fact, this skew also depends on how the clocks themselves are generated and on how large the relative delays are.

In this decimator, a common design technique is used to face the inter-clock skew. The method consists in adding a register at the interface between two clock domains. This additional register is clocked with an inverted version of the faster clock (Fig. 12.20). Such a structure is highly insensitive to the skew between $clk1$ and $clk2$. The demonstration of this statement is left to the reader.

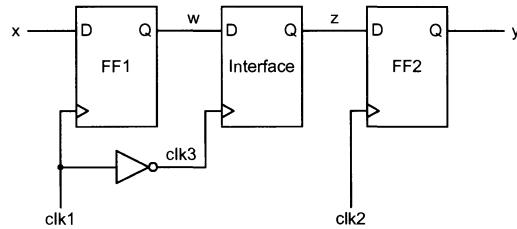


Figure 12.20: Example of an interface between the two clock domains $clk1$ and $clk2$. Here, $clk1$ is the faster one.

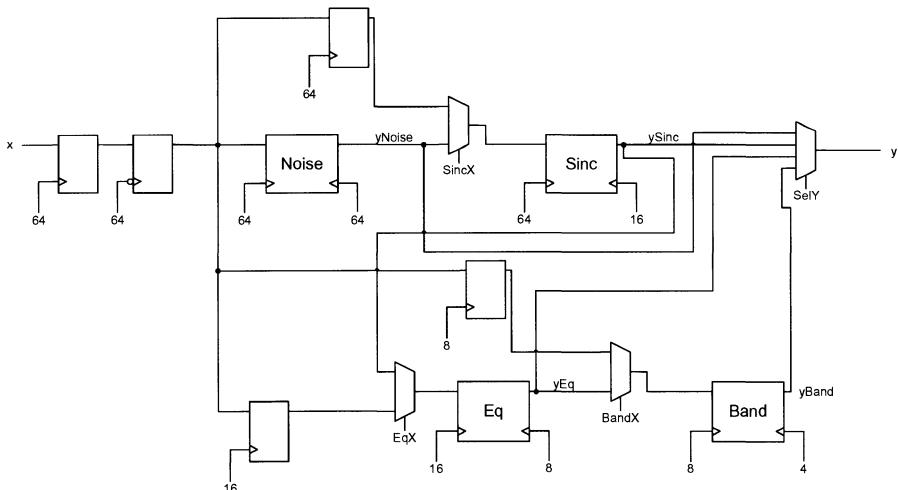


Figure 12.21: Block diagram of the decimation filter. The block “Noise” contains the noise cancellation logic.

The filter was provided with testing capabilities. The testing scheme adopted only requires few multiplexers and clock domain interfaces, while still allowing

the stages to be tested separately. By properly setting the four control signals, any set of adjacent stages can be tested. The exact details are omitted for the sake of brevity.

The block diagram of the decimation filter is depicted in Fig. 12.21. All the registers are intended to enforce the three-register clock domain interfacing scheme of Fig. 12.20. Whatever the combination of stages under test may be, when passing from a clock domain to another, the signal will always find the proper clock domain interface.

12.7.2 Sinc filter

This is the first stage of the decimation filter. Its decimation ratio and input bus width are 4 and 8, respectively.

The frequency response of a generic sinc^N filter in the z domain can be written as follows.

$$H(z) = \left(\frac{1 - z^{-M}}{1 - z^{-1}} \right)^N = \left(\sum_{k=0}^{M-1} z^{-k} \right)^N \quad (12.28)$$

If, like in our case, M is even, then the above expression becomes

$$H(z) = \left(\frac{(1 - z^{-M/2})(1 + z^{-M/2})}{1 - z^{-1}} \right)^N = (1 + z^{-M/2})^N \left(\sum_{k=0}^{\frac{M}{2}-1} z^{-k} \right)^N \quad (12.29)$$

When M is a power of two ($M = 2^P$), the process can be iterated, leading to the following result, which could be easily proved by induction:

$$\sum_{k=0}^{2^P-1} z^{-k} = \prod_{k=0}^{P-1} (1 + z^{-2^k}) \quad (12.30)$$

In our case, $M = 4$ and $N = 6$. Therefore, $H(z) = (1 + z^{-1})^6 (1 + z^{-2})^6$. This expression of the z -transform is directly implemented in the block diagram of Fig. 12.22.

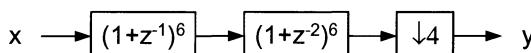


Figure 12.22: The non-recursive architecture for the sixth-order sinc filter.

By keeping in mind that the decimation switch can be partitioned in two cascaded blocks, each one with a decimation ratio of 2, and by applying the

commutative rule, the architecture in Fig. 12.23 is obtained.

In this way, the sinc stage is further partitioned into two identical sub-stages.

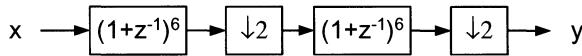


Figure 12.23: Efficient implementation of the non-recursive architecture.

The block diagram of the sinc filter is represented in Fig. 12.24. This is the direct implementation of the architecture of Fig. 12.23. The first register is a simple pipeline register and is not part of any clock domain interface. The two decimation switches are realized with two clock domain interfaces. As explained above, this is the safest way to discard samples. Due to the architecture chosen, the output bus size is 20.

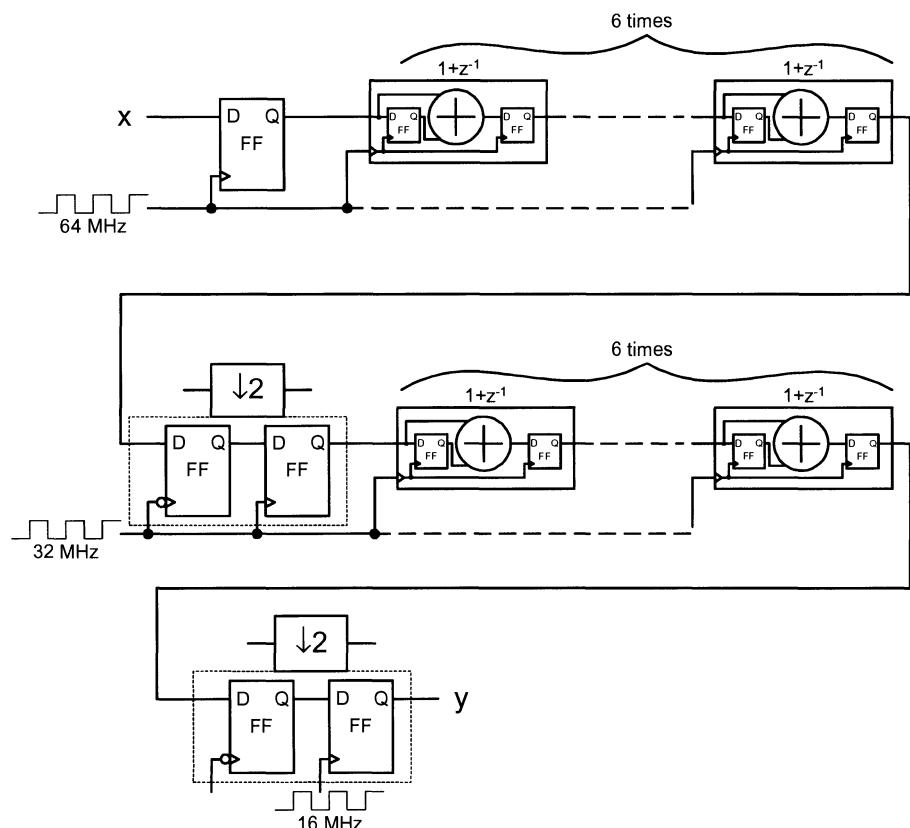


Figure 12.24: Block diagram of the sinc filter.

12.7.3 Equalizer

The equalizer has a decimation ratio of 2 and an input bus width of 16. Four bits are truncated from the output of the sinc stage. No multipliers are used and the CSD representation of the coefficients is exploited.

Two clock signals are sent to the stage from the top-level clock divider, namely one for the shift register and the other for all the remaining blocks. The frequency of the former is 16 MHz, while the latter has a frequency of 8 MHz.

Fig. 12.25 shows the block diagram of the equalizer. Like in the sinc stage, decimation is implicitly carried out by the clock domain interfaces.

This stage has 15 taps and is partitioned into five blocks. The clock interfacing circuitry is split into two parts and is embedded partly into the shift register and partly into the first block of adders.

12.7.4 Band filter

This stage is the last one and has a decimation ratio of 2. It was realized with 101 taps. The input bus width is 16, which means that 8 bits are truncated from the output of the equalizer. The same design considerations made for the equalizer also hold for the band filter.

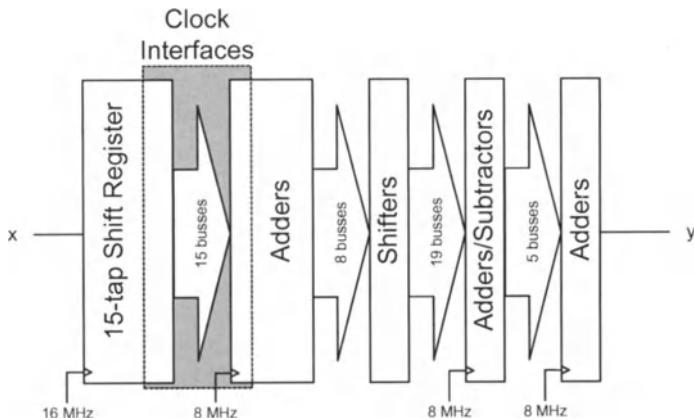


Figure 12.25: Block diagram of the equalizing filter.

12.8 REFERENCES

- [1] C. E. Shannon, “Communications in the presence of noise”, *Proc. IRE*, vol. 37, pp. 10-21, Jan. 1949.
- [2] D. A. Linden, “A discussion of sampling theorems”, *Proc. IRE*, vol. 47, pp. 1219-1226, July 1959.

- [3] R. E. Crochiere and L. R. Rabiner, "Interpolation and decimation of digital signals – A tutorial review", *Proc. IEEE*, vol. 69, pp. 300-331, March 1981.
- [4] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1983.
- [5] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial", *Proc. IEEE*, vol. 78, pp. 56-93, Jan. 1990.
- [6] T. Chen, L. Qiu, and E. W. Bai, "General multirate building structures with application to nonuniform filter banks", *IEEE Trans. Circuits and Systems – II: Analog and Digital Signal Processing*, vol. 45, pp. 948-958, Aug. 1998.
- [7] H. Johansson and L. Wanhammar, "Filter structures composed of all-pass and FIR filters for interpolation and decimation by a factor of two", *IEEE Trans. Circuits and Systems – II: Analog and Digital Signal Processing*, vol. 46, pp. 896-895, July 1999.
- [8] L. B. Jackson, *Signals, Systems, and Transforms*, Addison-Wesley, Reading, MA, USA, 1991.
- [9] A. V. Oppenheim and A. S. Willsky, *Signals & Systems (2nd ed.)*, Prentice-Hall, Upper Saddle River, NJ, USA, 1997.
- [10] S. Chu and C. S. Burrus, "Multirate filter designs using comb filters", *IEEE Trans. Circuits and Systems*, vol. 31, pp. 913-924, Nov. 1984.
- [11] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1975.
- [12] K. Azadet and C. J. Nicol, "Low-power equalizer architectures for high-speed modems", *IEEE Communications Mag.*, vol. 36, no. 10, pp. 118-126, Oct. 1998.
- [13] P. Pirsch, *Architectures for Digital Signal Processing*, John Wiley & Sons, Chichester, UK, 1998.
- [14] J. C. Candy, "Decimation for sigma delta modulation", *IEEE Trans. Communications*, vol. 34, pp. 72-76, Jan. 1986.
- [15] K. Y. Khoo, A. Kwentus, and A. N. Wilson Jr., "A programmable FIR digital filter using CSD coefficients", *IEEE J. Solid-State Circuits*, vol. 31, pp. 869-874, June 1996.
- [16] A. Chandrakasan, *Low-power digital CMOS design*, PhD thesis, University of California at Berkeley, UCB/ERL Memorandum No. M94/65, Aug. 1994.
- [17] A. P. Chandrakasan, M. Potkonjak, R. Mehra, J. Rabaey, and R. W. Brodersen, "Optimizing power using transformations", *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 14, pp. 12-31, Jan. 1995.
- [18] S. Devadas and S. Malik, "A survey of optimization techniques targeting low power VLSI circuits", *Proc. 32nd ACM/IEEE Design Automation Conference*, San Francisco, CA, USA, June 1995.

Chapter 13

LINE DRIVERS : EFFICIENCY, LINEARITY, BANDWIDTH

The Last Barrier in Wire-line Communications

Tim Piessens and Michiel Steyaert

Abstract In this chapter the last interfacing block to the line, being the line driver is addressed. The use of multi-tone modulation schemes in recent telecommunication systems raise the power consumption of these building blocks enormously, creating a new bottleneck. Soon, the classical class AB solutions will be no longer applicable from a power efficiency point-of-view. More advanced circuit techniques like class G and class H are emerging. Ultimately switching type amplifiers, which suffer the high bandwidth and linearity constraints, will have to be used to decrease the power dissipation.

13.1 INTRODUCTION

Due to the rising need for broadband communication and broadband connections to the Internet, the research on power amplifiers increased significantly. Novel modulation schemes like Discrete Multi-Tone modulation (DMT) and Orthogonal Frequency-Division Multiplexing (OFDM) are developed in order to fully deploy lossy channels like the twisted pair copper wires used in the xDSL application fields [1–3]. The major disadvantage of this technique is the high Crest Factor ($CF = \frac{V_{peak}}{V_{RMS}}$) of the time domain signal associated with these kind of modulations. Since the efficiency of linear type power amplifiers is inverse proportional to the CF, power efficiency is the major problem for these kind of modulation schemes. In this chapter mainly the line driver problems in Digital Subscriber Loop (xDSL) applications are addressed since this seems to be the present driving application in this research field.

Figure 13.1 shows the comparison of the relative area and power consumption of two typical commercial available Asymmetrical Digital Subscriber Loop (ADSL) chip-sets. For this the values of different chip factories were taken and a relative mean has been calculated. In the right of figure 13.1 the same calculation was redone

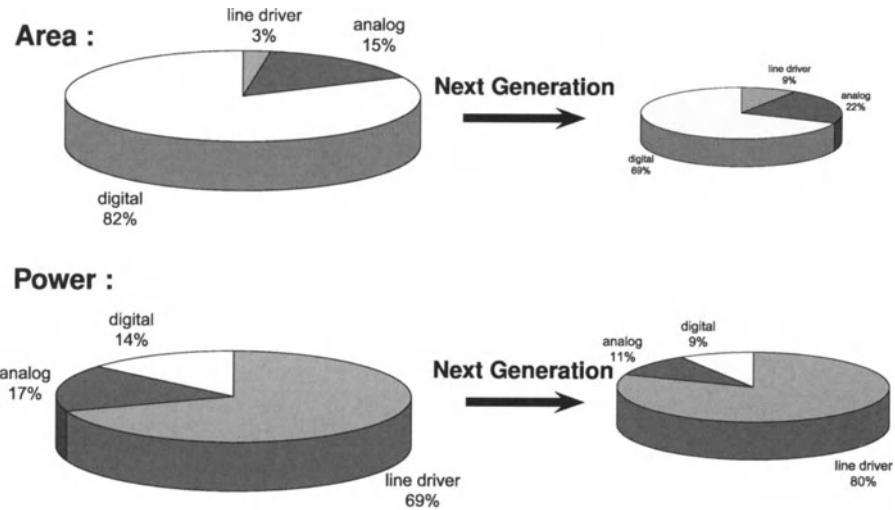


Figure 13.1: Relative portion of the area and power consumption budget for different building blocks of a commercial Central Office (CO) chip-set and its evolution towards a next generation.

but then for the consecutive chip-sets of the same foundries. One can clearly notice the shrinking effect between two consecutive chip-sets. This is mainly due to the smaller process technology used and the digital shrinking logically following this advance. From a power consumption point-of-view however the decrease is not that significant. A closer inspection reveals the line driver as being the real bottleneck. In the most recent generation of xDSL chip-sets the relative portion of the line drivers power consumption has increased to 80% of the total power budget.

In the second section this phenomenon will be further explained by looking further into the xDSL modulation schemes. The basic properties of a DMT modulated signal and the requirements for a working xDSL system will be combined to derive some basic limits for the design of an xDSL line driver. In section 13.3 the linear type amplifiers class AB will be thoroughly discussed. These type of amplifiers are still the most commonly used line drivers for their low distortion characteristics. In order to save power these amplifiers are evolving towards linear amplifiers with modulated supplies (classes G and H). However in order to reach a sub 500 mW consuming line driver, switching type amplifiers should be investigated. In section 13.4, these kind of amplifiers are further explored. A special case, the Self-Oscillating Power Amplifier (SOPA) will be the subject of a separate section (section 13.4.2).

13.2 WIRE-LINE COMMUNICATION

13.2.1 xDSL-Technologies

In the market to provide broadband Internet access to the home, the xDSL family has proven to be a valiant alternative. In this text we will focus mainly on the DMT-based variants ADSL and Very high-speed Digital Subscriber Loop (VDSL) since they are the most demanding technologies for the line drivers [4].

In table 13.1 the most important properties for line driver design of ADSL and VDSL are summarized. To compare it with older technologies one can clearly notice the enormous shift in line driver requirements for the two modem types. In less than one decade of xDSL research activities, an improvement of over 3 decades in bit-rate has been reached. This however at a drastic cost in the Analog Front-End (AFE) design specification. The linearity specifications are not significantly relaxed, but the bandwidth and necessary output power are increased towards the current technology limits. The specific linearity specifications Missing Tone Power Ratio (MTPR) and Missing Band Depth (MBD) will be discussed in subsection 13.2.2 where the DMT modulation will be discussed in more depth.

The requirement for a minimal power dissipation is a direct consequence of the deployment scenarios of xDSL communication. In an xDSL system, the Customer Premises Equipment (CPE, = equipment at the users side) is with a single twisted

Table 13.1: Summary of the most important xDSL requirements for line driver design. The V.34 standard is added in grey for comparison reasons.

	<i>Distortion</i>	<i>Bandwidth</i>	<i>Output Power</i>
V.34	THD <-70 dB	4.96 kHz	0 dBm
ADSL(-Lite) US		103.5 kHz	13 dBm
ADSL-Lite DS	MTPR > 34 dB	418.3 kHz	16.3 dBm
ADSL DS	MTPR > 55 dB	970.3 kHz	20 dBm
VDSL DS	MBD > 63 dB	8.5 MHz	14.5 dBm

DS = downstream, US = upstream

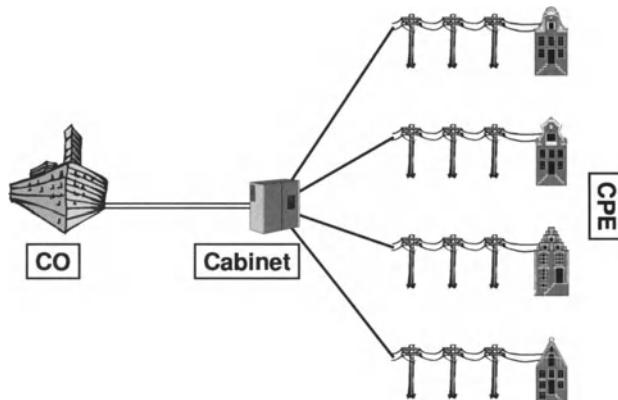


Figure 13.2: Typical deployment of a VDSL system from a cabinet case.

copper wire-pair connected to the Central Office (CO) . Since the attenuation of a copper wire is the limiting factor for the achievable bit-rate, the loop-lengths have to be kept short. For customers who reside too far from the telephony central, the CO-modem has to be deployed from a street-side cabinet case. This is mostly the case for VDSL-modems since loop lengths are becoming very small. This is depicted in figure 13.2. These cabinet cases don't have temperature control except natural heat convection which is limited due to its small volume. The maximum operating temperature and thus maximum heat dissipation limits the amount of installable lines and thus the number of servable customers.

The stringent linearity requirements are a direct consequence from the fact that the twisted copper wires are not shielded. Therefor the out-of-band specifications of the xDSL modem are very stringent to avoid self induced cross-talk and e-gress radiation into other frequency bands like the ones used for amateur radio.

For the choice of a process technology, one has to take into account that the twisted pair is also shared with the Plain Old Telephony System (POTS) . On the copper wire high voltages occur due to the POTS system. A line interface that is able to handle POTS and xDSL on the same die, has to take these high voltages into account [5–7].

As a summary, due to the high performance of an xDSL system, the line driver has to :

- Drive a relative large power to the line
- with a high linearity
- in a high bandwidth.
- The power dissipation needs to be minimal to allow maximal installable lines at the CO-side
- Out-of-band specifications are very stringent (<-100 dBm/Hz)
- Deal with high voltages

13.2.2 Basic Properties of Multi-Tone Channels

Figure 13.3 shows a schematic overview of an ADSL system. The Discrete Multi-Tone (DMT) modulation in this case consists of 22 carriers containing the upstream (=from CPE to CO) information and 221 downstream (=from CO to CPE) carriers. Each carrier is Quadrature Amplitude Modulation (QAM) modulated. The constellation size of each carrier is given by the signal-to-noise ratio (SNR) of the channel at the specific carrier frequency [8]. The time domain representation of the signal is given in figure 13.4. The ADSL downstream signal can be represented by

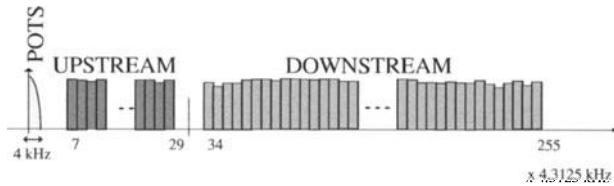


Figure 13.3: Spectrum of an ADSL system.

equation 13.1

$$x(t) = \sum_{n=34}^{255} a_n s(t - nT) \cos(2\pi f_i t) + \sum_{n=34}^{255} b_n s(t - nT) \sin(2\pi f_i t) \quad (13.1)$$

In equation 13.1, a_n and b_n represent the constellation point of the n-bit QAM modulation, $s(t)$ is an envelope function used to normalize the average energy per channel. The carrier spacing f_i is set to 4.3125 kHz for the ADSL system. Since the constellation points can be regarded as random variables for normal data communication and the number of carriers is sufficiently large for the central limit theorem to hold [9], the amplitude distribution f_A of the DMT-signal $x(t)$ can be approximated as being Gaussian (13.2).

$$f_A = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{A^2}{2\sigma^2}\right) \quad (13.2)$$

σ is the rms-voltage of the signal. From equation 13.2, one can calculate the distribution function for the CF in an average DMT generated symbol with N

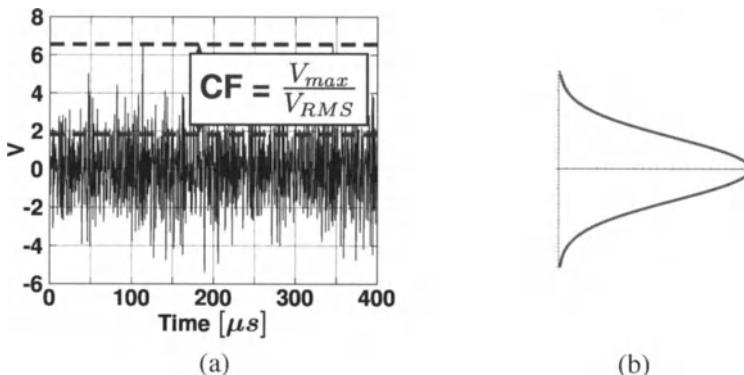


Figure 13.4: Time domain representation of a DMT-modulated signal (a) and amplitude distribution (b).

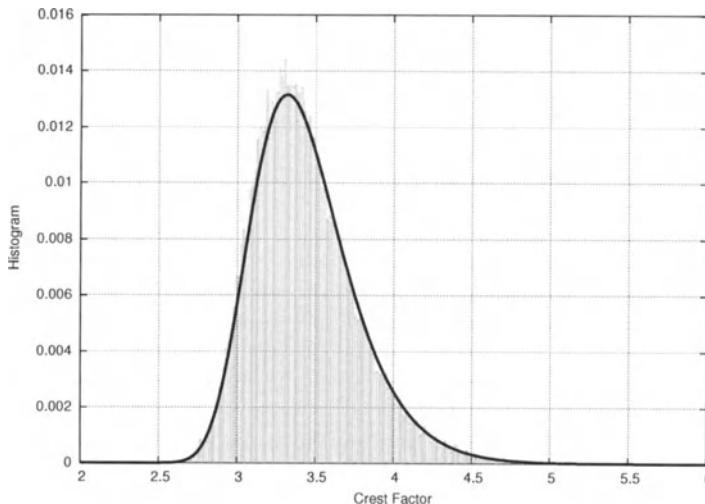


Figure 13.5: Distribution function for the crest factor of an upstream ADSL symbol (solid line) compared with a monte carlo simulation (grey histogram).

carriers.

$$f_{CF} = \frac{2\sqrt{2}}{\sigma\sqrt{\pi}} N \left(\operatorname{erf} \left(\frac{\sqrt{2}CF}{2\sigma} \right) \right)^{2N-1} \exp \left(-\frac{CF^2}{2\sigma^2} \right) \quad (13.3)$$

From equation 13.3 the crest factor distribution for an ADSL upstream signal can be calculated. The result is plotted in figure 13.5 and compared with the histogram obtained from a monte carlo simulation. Herefor, the histogram of the crest factor of 50000 randomly generated 4QAM modulated ADSL downstream signals has been calculated. This Fisher-Tippet distribution has a slow tail when going to high crest factors. Unlike the ideal Fisher-Tippet distribution, the distribution of the crest factors of an ADSL-downstream signal has a maximum of 22.6, when all the carriers are in phase.

In practical designs the crest factor has to be limited. The clipping of the CF will lower the quantization noise of the used Analog-to-Digital Converter (ADC) and the Digital-to-Analog Converter (DAC) and their power consumption, but will introduce clipping noise [9]. In practical ADSL systems the CF is limited to a factor of 15 dB or 5.6.

Due to the stochastic nature of the time-domain signal, a non-linearity will generate a rise in the noise-floor of the complete signal. This ‘distortion-noise’ can be calculated with the cross-correlation of the original signal with the output of the

non-linear characteristic $g(x)$ [10].

$$\sigma_d^2 = \int_{-\infty}^{\infty} (x - g(x))^2 \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \quad (13.4)$$

Clipping noise can be regarded as a special case of (13.4).

$$\sigma_d^2 = \int_A^{\infty} (x - A)^2 \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \quad (13.5)$$

Due to the stochastic nature of the DMT-signal, a direct relation between Signal-to-Noise Ratio (SNR) and MTPR cannot be calculated without information on the nature of the non-linearity.

This distortion-noise will not only be the limiting factor for the in-band bit-rate, but the out-of-band distortion-noise will jam the upstream signal coming from the CPE-side. This distorted echo-signal is very hard to compensate in the digital domain.

13.3 LINEAR POWER AMPLIFIERS

Due to the high bandwidth and the high linearity specifications of an xDSL system, a linear power amplifier is the natural choice. In this type of amplifier the active elements will conduct the output currents while there is an output voltage over these elements. The efficiency of this type amplifiers will thus be very low.

13.3.1 Class AB

13.3.1.1 Class AB operation

The class AB operation of a power amplifier is an intermediate between class A and class B. Class A is the most linear line driver. This is due to the fact that the active element (output driving transistor) is never switched off. This comes with a significant heat-dissipation and thus very low efficiency. Class A operation is depicted at the left of figure 13.6. Above the input-output relationship in figure 13.6 the efficiency for a sinusoidal signal is depicted. For a crest factor of 5.6 and a rail-to-rail output driving stage (which is in practice impossible) the maximum efficiency for a sinusoidal signal will be lower than 9%. The theoretical minimal power dissipation of a class A power amplifier for ADSL is more than 2 W, so the class A operation will be completely unacceptable for an xDSL system.

The class B operation gives a higher efficiency by taking two active elements in a push-pull configuration. Power is saved since the pull transistor can be turned off for the rising edge and vice-versa. For a DMT-signal the efficiency can be calculated

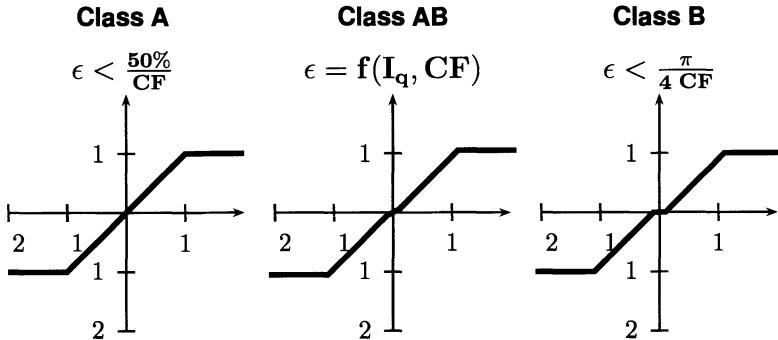


Figure 13.6: Input-output characteristics for class A (left) and class AB (middle) and class B (right) operation.

as in equation (13.6).

$$\epsilon = \frac{\sqrt{2\pi}}{2} \frac{V_{rms}}{V_{DD}} = \frac{\sqrt{2\pi}}{2} \frac{1}{CF} \frac{V_{swing}}{V_{DD}} \quad (13.6)$$

If we assume a rail-to-rail output swing a maximum efficiency of 22% can be reached. For a resistive terminated ADSL-line, the theoretical class B line driver thus dissipates 710 mW.

However due to the threshold voltage (V_T) of a transistor, the two active elements of the push-pull stage are both turned off around the zero-crossing. Therefor the input-output characteristic of an ideal class B power amplifier shows a dead zone. Considering the general distortion-noise calculation (13.4), the distortion will be active in the region where the signal resides most of the time. The distortion-noise associated with crossover-distortion can be calculated as follows

$$\sigma_d^2 = \int_{-V_T}^{V_T} x^2 \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \quad (13.7)$$

To cope with the cross-over distortion, a biasing scheme between the two active elements is added. In this way the elements are never turned off. At the zero-crossing, a small current, the quiescent current, flows through both the active elements. The dead zone is thus removed from the input-output characteristic. Equation 13.7 has to be altered thus with the quiescent current linearized input-output characteristic

$g(x, I_q)$.

$$\sigma_d^2 = \int_{-V_T}^{V_T} (x - g(x, I_q))^2 \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \quad (13.8)$$

The quiescent current (I_q) has to be accurately controlled, since an I_q which is too low, will generate too much distortion. The higher the I_q , the closer the amplifier will get to class A operation, thus the lower the efficiency. For the calculation of the overall efficiency the power-consumption of the quiescent current control circuitry has to be added (P_{qcc}).

$$\epsilon = \frac{V_{rms}^2 \sqrt{\pi}}{V_{DD}(V_{rms}\sqrt{2} + I_q\sqrt{\pi}) + P_{qcc}\sqrt{\pi}} \quad (13.9)$$

Assuming a rail-to-rail output and a fixed CF for DMT-modulation, the efficiency for a class AB line driver can be written as

$$\epsilon = \frac{V_{DD}^2 \sqrt{\pi}}{CF V_{DD}(V_{DD}\sqrt{2} + CF I_q\sqrt{\pi}) + CF P_{qcc}\sqrt{\pi}} \quad (13.10)$$

The efficiency for a class AB line driver increases with increasing supply voltage [5], ultimately approaching the class B limit. This however requires a more expensive, high frequent, high voltage technology. The class AB operation is currently the favorite xDSL line driver [6, 11–13], for its high linearity.

13.3.1.2 Quiescent current control

When designing class AB line drivers, the quiescent current control block is the most critical block for the performance of a class AB line driver.

Figure 13.7 shows the dependence of the quiescent current and the transistor sizing on the signal to total harmonic distortion ratio, taken from simulation results presented in [14]. In this simulation three distinct regions can be observed. In the low I_q regions, the cross-over distortion is dominant in the signal to distortion ratio. The obtained Signal-to-Noise-and-Distortion (SNDR) levels are independent from the relative transistor sizing. In the middle region, the class AB line driver is in its optimal performance region. SNDR becomes highly dependent on the transistor sizing since the maximal driving capacity defines the output signal level. In the high I_q region, the output transistors are put more quickly in the linear operating region. The distortion level rises due to signal clipping.

A first approach to lower the influences of the cross-over distortion has been presented in [15]. Two error amplifiers are placed in connection with a pseudo push-pull connected pair to lower the cross-over distortion. This is schematically presented in figure 13.8. The increase in SNDR can be easily calculated (13.11).

$$HD_{CL} = \frac{HD_0}{A_{preamp} A_{EP|N} gm_{M1a|b} R_{load}} \quad (13.11)$$

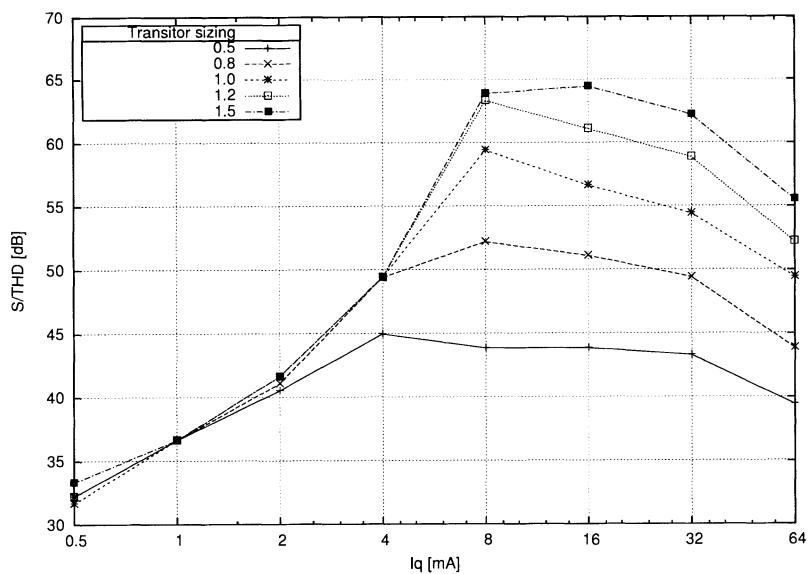


Figure 13.7: Signal to total distortion ratio versus quiescent current [14].

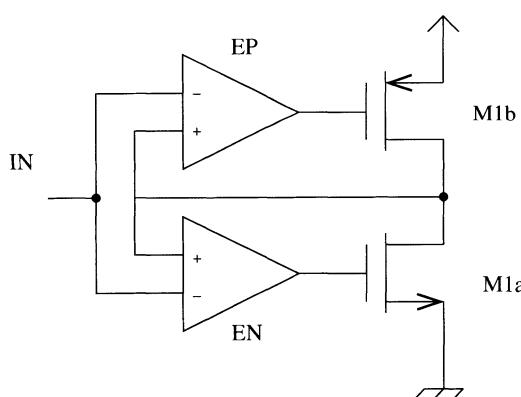


Figure 13.8: Line driver with included error-amplifiers to lower cross-over distortion.

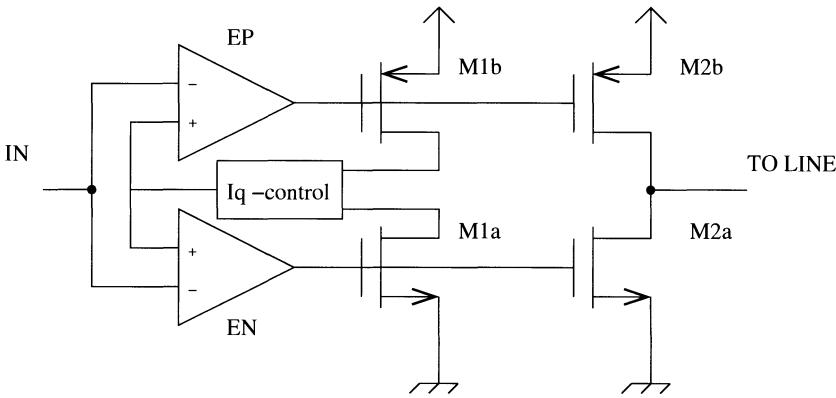


Figure 13.9: Quiescent current control technique to overcome limited error-amplification due to random offset.

So the initial amount of output-related cross-over distortion HD_0 is lowered by the loop gain containing the error amplifier gain and the gain of the pre-amplifier. This reduction however will be limited by the offset voltage of the error amplifiers. The estimated maximum variation on the quiescent current ΔI_q can be calculated as

$$\frac{\Delta I_q}{I_q} = \frac{2V_{offset}A_{EP}}{(V_{gs} - V_T)_{M1a|b}} \quad (13.12)$$

Together with the results from figure 13.7 and using a technology with reasonable matching, the maximum error amplification A_{EP} is around 8.

To overcome the problem of a random offset in the error-amplifiers, a quiescent current control scheme has to be adopted. Figure 13.9 shows schematically an approach adopted in [14]. The output transistors are decoupled in order to measure the quiescent current through the transistors $M_{1a|b}$. Since the output transistors are quite large this measurement will match the quiescent current through the actual output transistors $M_{2a|b}$. The measured I_q will then be compared with a reference current I_{ref} . The quiescent current control circuit used in [14] consist of two Schmitt triggers (one for the output PMOS, one for the output NMOS) which trigger a charge pump. This charge pump together with the dump capacitor form a filter with a very small time constant in order to only get the DC quiescent current. The obtained error signal is added at the inputs of the error amplifiers to counteract the random offset. An SNDR of 62dB for a 100 mW signal has been reported using this technique. The reference current can be generated on-chip by a bandgap reference. This however limits the flexibility of the control loop.

Controllability has been gained in [12] by replacing the analog quiescent control loop by a digital one. The cost to be paid is a loss in efficiency by the integration of a measurement ADC and a controlling DAC.

13.3.1.3 Final Remarks on class AB

The class AB line driver is still the most widely used line driver for its high intrinsic linearity and its well known behavior. New circuit techniques allow designers to construct quiescent current control loops which have a lower power consumption but provide higher linearities over a wider mismatch range.

However due to its very low efficiency, the class AB line driver is the power bottleneck for xDSL systems and will become unusable in future central office applications.

13.3.2 Class G,H

The class G line driver is a logical solution for the efficiency problem related with class AB line drivers. The efficiency of a class AB line driver is inverse proportional with the ratio of the rms voltage V_{rms} and the supply voltage V_{DD} (13.9). In a class G design, multiple supplies are connected to a class AB line driver. In this section we will further focus on the case were two supplies are used : one scaled supply for the main signal V_{DD}/n and one higher for the peak amplitudes V_{DD} . Figure

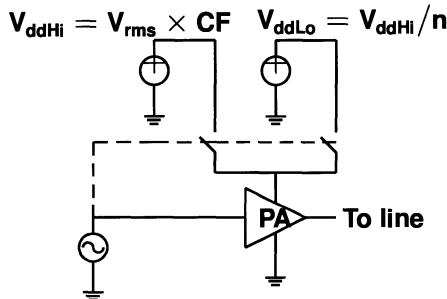


Figure 13.10: Principle scheme of a 2 supply class G power amplifier.

13.10 shows the topology which is mostly used nowadays [16]. The efficiency for an ideal two-supply class G amplifier can be calculated as follows :

$$\epsilon = \frac{n V_{rms} \sqrt{\pi}}{V_{DD} \sqrt{2} \left(\left(1 - \exp \left(\frac{-V_{DD}^2}{2n^2 V_{rms}^2} \right) \right) + n \exp \left(\frac{-V_{DD}^2}{2n^2 V_{rms}^2} \right) \right)} \quad (13.13)$$

Figure 13.11 shows the obtained efficiency for different values of $n = \frac{V_{ddHi}}{V_{ddLo}}$ with

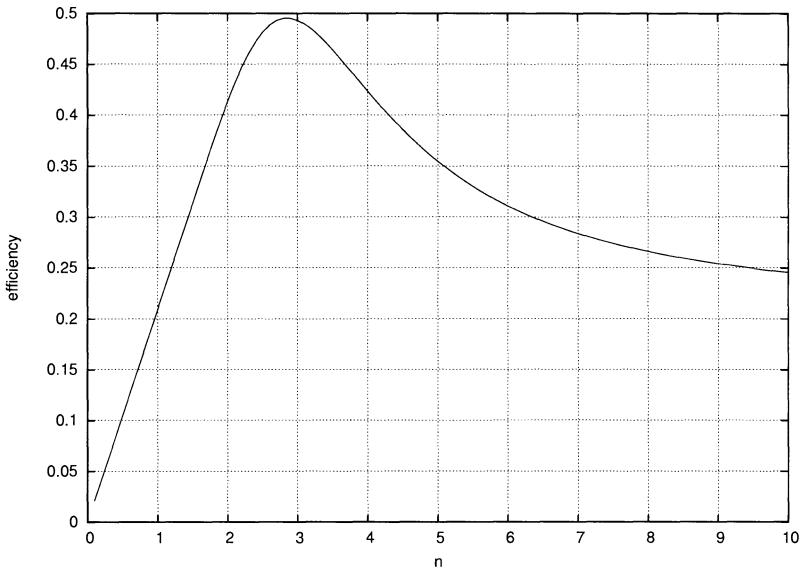


Figure 13.11: Efficiency versus V_{DDHi}/V_{DDLow} ratio for an ADSL input signal.

V_{DDHi} obtained from the crest factor constraints described in section 13.2.2. Values below $n=1$ create supply voltages that are too high for the crest factor constraint, resulting in an extreme low efficiency. At $n=1$ the class AB efficiency is obtained. The efficiency approaches a maximum around $n=3$. This maximum is below 50% for a 2 supply class G amplifier. For higher values of n more amplitudes are amplified using the high supply. Ultimately the class AB efficiency is reached.

These figures are however to optimistic and state-of-the-art class G xDSL power amplifiers show efficiencies close to their high-voltage class AB counterparts [16]. This is mostly due since :

- The matching between the delays of the forward signal path through the class AB power amplifier and the supply steering path (dashed line in figure 13.10) is of the utmost importance. Any mismatch will generate clipping noise (13.5) like a DMT-signal clipped by $V_{DDLO} = V_{DD}/n$. Since this timing delay constraint is relative to the bandwidth of the signal, it will be even more important towards the faster xDSL types. Therefor class G will only be feasible when the detection circuitry, the class AB power amplifier and the supply switches are integrated on the same die. To relax the matching constraint, a larger envelope around the voltage peak has to be generated.

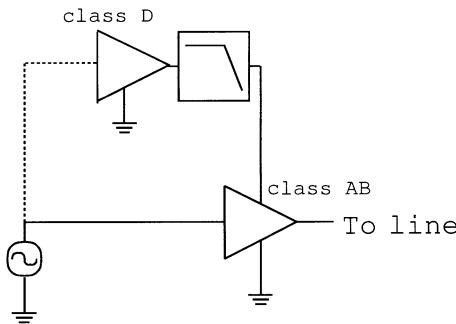


Figure 13.12: Principle scheme of the class H power amplifier.

This comes at the cost of a more elaborate analog signal processing and less efficiency due to a suboptimal supply switching scheme.

- The generation of the switching signal is very critical. The detection of the input signal at the trigger level for the switches has to be accurate enough to avoid unnecessary supply switches. A comparator with hysteresis is mostly used to gain noise-immunity. However if we take this hysteresis into account in formula 13.13, a hysteresis of 2.5% will already cause a power efficiency drop of 5%.
- The extra supply voltages need to be generated as well. So the efficiency of the DC-DC converter has to be added to the efficiency formula (13.13). State-of-the-art DC-DC converters tend to reach an 80% efficiency [17]. If the high voltage is generated a 3% efficiency drop can be calculated. For the generation of the low voltage, this decay will become 6%.

Due to these constraints it is believed that adding more power supplies to a DMT modulated signal class G power amplifier will not be beneficial from a power efficiency point of view.

Figure 13.12 shows the principle schematic of the class H amplifier. The supply voltage of the class AB line driver in the class H configuration is modulated directly by the input voltage. It can also be regarded as a class G power amplifier with an infinite amount of supply voltage sources.

The problems of the class AB line driver shift now to the supply driver. Therefor the class H approach is only useful when the supply source is constructed as a switching type line driver. The class AB line driver acts then as a linearizing power amplifier that linearizes the intrinsic non-linearity of the class D supply driver (see also 13.4.1). This technique is heavily used in high-performance, high efficiency

audio power amplifiers [18], for xDSL line drivers however the class H is still beyond the state-of-the-art since :

- The class D line driver that has to be used to drive the supply voltage needs to be a very high frequency switching type line driver. The mean switching frequency of this line driver has to be filtered out, so the switching does not turn off the forward class AB line driver. Passive filtering is needed between the class D power driver and the supply connection of the class AB (see fig. 13.12). This passive filter will be bulky and power consuming, lowering the overall efficiency.
- The bandwidth and linearity of the class D line driver has to be as stringent as for the direct steering with a switching type line driver, which will be more elaborated in section 13.4. The voltage spikes in a DMT-signal are constructed by the in-phase component of several carriers. Any non-linearity that deteriorates the phase-characteristic of the DMT-signal will generate the amplitude spike at the wrong time instant.
- The timing-delay mismatch problem between the forward class AB line driver and the class D power driver is comparable with the class G integration problems. Analog signal preprocessing is thus necessary to generate a close but accurate envelope.
- The integration of a switching type line driver with a linear one on the same die will generate substrate-noise related issues. These have to be taken into account not only in the design phase but also in the layout of the complete system.

13.4 SWITCHING TYPE LINE DRIVERS

Switching type line drivers have the intrinsic possibility to reach a 100% efficiency, since there is no power dissipation in an ideal switch. In this section we will discuss two types of power amplifiers. First the synchronous class D type (also called class S) line driver will be discussed. The mean switching frequency is related to a central clock frequency. A novel type asynchronous Self-Oscillating Power Amplifier (SOPA) will be the subject of a second section.

13.4.1 Class D

Basically the class D power amplifier architecture is constructed as depicted in figure 13.13. It consists of three major building blocks :

- 1.- A Pulse Width Modulator (PWM) that is driven by an external clocking signal. This modulator will convert the input signal in a switched signal that is able

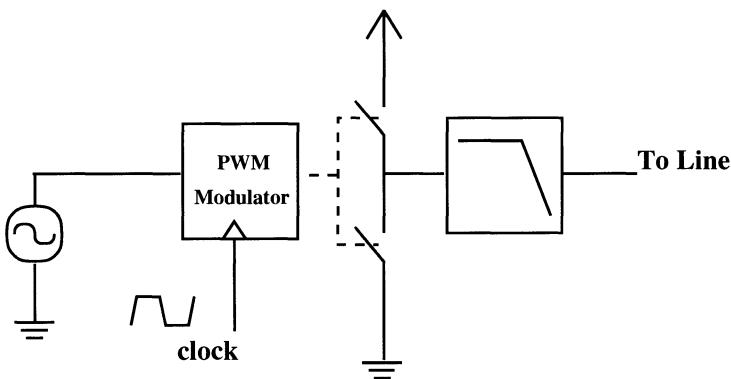


Figure 13.13: Principle schematic of a class D power amplifier.

to drive the output switches. The mean switching frequency is determined by an external clocking signal. Later in this section some modulation schemes will be further elaborated.

- 2.- The output-switches need to convert the small-power switching signal in a high-power switching signal. Their sizing will directly influence the complete drivers performance. Too small switches have an on-resistance that is too large, degrading the overall efficiency. Switches that are too large will become too slow. This will generate distortion since the amount of charge that is pumped into the load will decrease due to larger rise -and fall-times. The significance of this charge-loss is dependent on the pulse width, thus on the input signal. Too slow switching will also cause a significant dissipation in the switches, degrading the efficiency.
- 3.- The output driver is followed by a passive filter to filter out the mean switching frequency. Since the out-of-band specifications of an xDSL system are very severe to avoid echo and e-gress, the mean switching has to be suppressed below the -100 dBm/Hz noise floor. The parasitic resistances connected with these type of filters however will lower the efficiency. Another disadvantage is that such a high-performance passive filter will consume a considerable amount of board-space.

The mean switching frequency, and thus the clock frequency, needs to be much higher than the signal bandwidth for the following reasons:

- The order of the filter is inverse proportional with the Over-Switching Ratio ($OSR = \text{mean switching frequency} / \text{signal bandwidth}$). Since the volume of

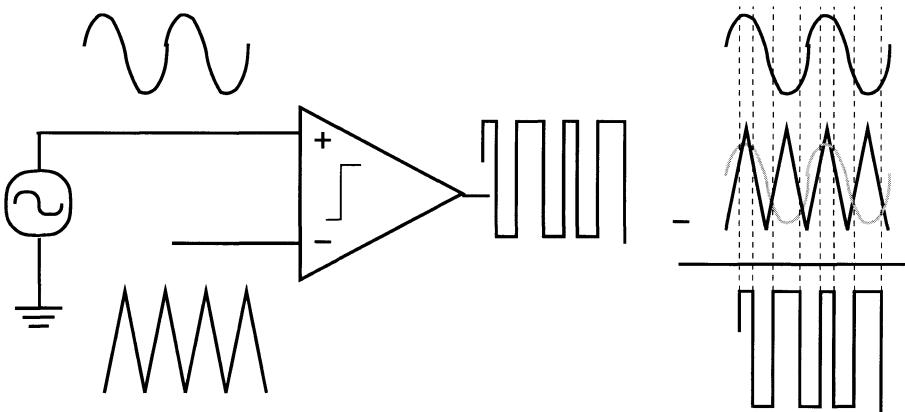


Figure 13.14: Basic PWM modulator block.

the filter and the parasitic resistance is proportional with the filters order, a low order filter and thus a high OSR will be most beneficial.

- Depending on the modulation scheme, the OSR has a minimum value in order to be able to generate a PWM signal with sufficient low distortion.

A basic PWM modulator building block is depicted in figure 13.14. This modulation scheme has the advantage of being very simple to implement and to be a very straightforward implementation of a PWM principle. The ideal modulation creates spurs at harmonic offsets from the switching frequency and a reasonable amount of energy at the clocking frequency. The clock frequency should thus be taken high enough. Another disadvantage of this topology [19] is its high sensitivity to circuit non-idealities:

- If the system depicted in figure 13.14 has to be linearized to achieve the desired performance, one needs to construct a feedback loop using a pre-amplifier in front of the comparator. The feedback signal than has to be taken behind the filter since the switching signal will take the pre-amplifier out of its linear working region, and to compensate the non-linearities of the high power output filter.
- The output filter cannot be made sharp enough to filter out the switching component completely, since the parasitic resistance limit the filter order. This remaining switching component will be amplified by the pre-amplifier and will arrive at the comparator were it most certainly will be out-phased

with the original triangular wave. The intermixing products will generate in-band distortion.

- High frequent ripple picked up from the environment at the inputs of the comparator can cause erroneous switching of the class D amplifier which cannot be corrected by feedback, since it is filtered out in the correcting loop. The comparator and pre-amplifier therefor have to be shielded very carefully.
- Pulse amplitude errors are due to changes in the output load. These changes are very common in wire-line communications.

An other modulation technique proposed to create a PWM modulated signal is based on $\Delta\Sigma$ modulation [20]. In this technique the errors are shaped towards higher frequencies. An external filter needs to filter out these high frequency errors. The major advantages of these techniques are

- The accuracy of the analog components is largely relaxed, compared with the PWM modulator.
- Since the high frequency errors are randomized, the $\Delta\Sigma$ modulation generates less in-band substrate noise and thus less self-interference.
- Baseband feedback is inherent in the system, so in-band distortion is lower for the same clocking frequencies.

Formula 13.14 calculates the Dynamic Range (DR) of a $\Delta\Sigma$ based switching power amplifier.

$$DR = \frac{3\pi}{2}(2n + 1) \left(\frac{OSR}{\pi} \right)^{2n-1} \quad (13.14)$$

The $\Delta\Sigma$ -loop has order n and an over-switching ratio (OSR). The order should be kept as low as possible, since the order of the output filter needs to be at least one order higher than the $\Delta\Sigma$ order. So in order to meet the necessary linearity specifications, a higher OSR has to be chosen. This however, requires very high frequent high-power devices, which are not available in present semi-conductor technologies.

13.4.2 Self Oscillating Power Amplifier

13.4.2.1 0th order SOPA

To overcome the problems related with clocked switching type line drivers a novel type of asynchronous line driver has been introduced [21]. The basic system schematic of this 0th order Self Oscillating Power Amplifier (SOPA) is depicted

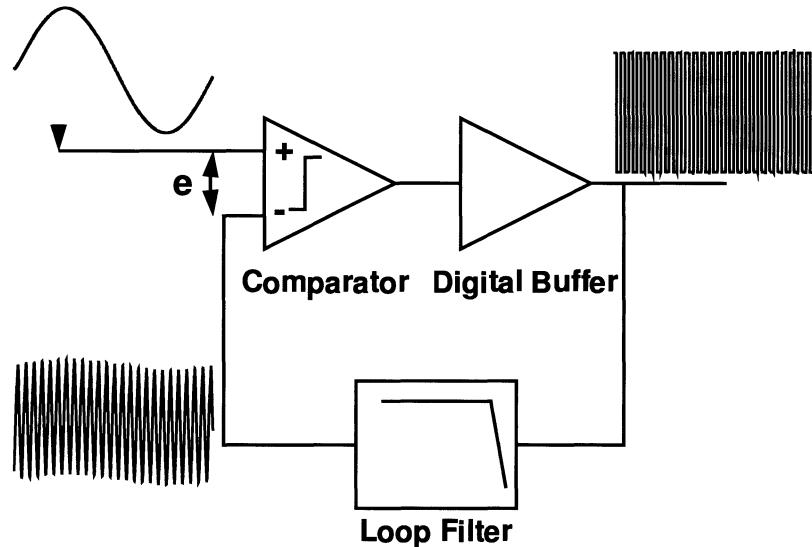


Figure 13.15: Principle schematic of a 0th order SOPA with characteristic waveforms.

in figure 13.15. The basic SOPA amplifier consists of three building blocks : a comparator, a digital driver and a loop-filter. Since the output of the SOPA is a switching signal, high efficiencies like in class D line drivers can be expected. No circuit element is clocked, so the loop is fully analog. The loop has been constructed to be is unstable. If no input-signal is applied, a limit cycle oscillation occurs in the loop. The limit cycle oscillation will dither the loop for any incoming signal. Therefor the limit cycle frequency and amplitude at the input of the comparator will be important design parameters.

Due to the filtering of the higher frequencies by the loop filter, the describing function analysis technique [22] can be used to analyze the hard non-linearity. The comparator with a dominant pole p_1 can be modeled in the frequency domain as the following single input describing function.

$$N(A) = \frac{2V_{DD}}{\pi A} \frac{p_1}{s + p_1} \quad (13.15)$$

The term A represents the limit cycles amplitude at the input of the comparator. The criterion for self oscillation with the loop filter $L(s)$ is:

$$1 + N(A)L(s) = 0 \quad (13.16)$$

Solving this complex equation gives us an expression for the limit cycle amplitude and frequency. The transfer function of the linearized system for error signals with amplitude e is given by

$$T(s) = \frac{N_2(A, e)}{1 + N_2(A, e)L(s)} \quad (13.17)$$

In which the dual-input describing function $N_2(A, e)$ describes the frequency domain input-output relation of the comparator in presence of a limit cycle oscillation with amplitude A for the forced signal with a given frequency and an amplitude at the comparator input e . In the closed loop SOPA system e will be the error signal. The describing function $N_2(A, e)$ can be calculated as follows:

$$N_2(A, e) = \begin{cases} \frac{V_{DD}}{\pi e} \left(\frac{e}{A} \right) {}_2F_1\left(\frac{1}{2}, \frac{1}{2}; 2; \left(\frac{e}{A}\right)^2\right) & \text{for } 0 < e < A \\ \frac{2V_{DD}}{\pi e} {}_2F_1\left(\frac{1}{2}, -\frac{1}{2}; 1; \left(\frac{A}{e}\right)^2\right) & \text{for } 0 < A < e \\ \approx N(A)/2 & \text{when } 0 < e \ll A \end{cases} \quad (13.18)$$

From equation 13.17 one can draw the following qualitative conclusions: when the error signal is reduced by the feedback loop in such a way that the error signal amplitude e becomes smaller than the limit cycle amplitude A calculated from (13.16), the transfer function becomes independent on the error signal itself. Thus

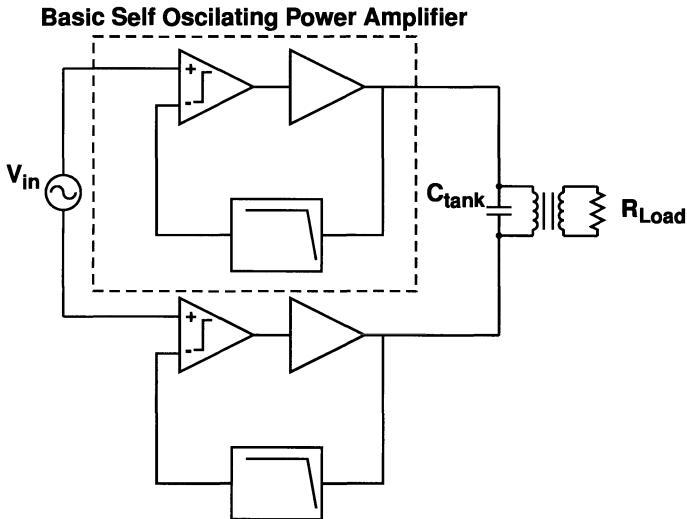


Figure 13.16: Full SOPA line driver coupled using a line transformer.

the switching comparator behaves as a linear amplifier for the input signal. Equation (13.17) in that way imposes also a limit on the maximum output signal swing of the system.

These conclusions can be easily understood from the waveforms in figure 13.15. At the negative input of the comparator the filtered version of the switching signal resides. This triangular wave with amplitude A will be shifted by the input signal. Since the signal bandwidth resides within the low-pass band of the loop filter, the system is in unity feedback. The resulting waveform at the negative input thus will be an oversampled triangular wave approximation of the input-signal. Due to the feedback, the frequency content in the pass-band of the loop filter and thus also in the signal bandwidth of the switched output signal will be the same as that of the input. Since the modulated signal is a discrete amplitude, continuous time signal, the quantization noise will be much lower than that of a class D PWM signal, which is discrete amplitude, discrete time, and this for lower over-switching ratios (OSR).

A complete SOPA line driver is a bridge-like connection of two basic SOPA amplifiers. The line transformer, which is always present in wire-line communications for galvanic isolation, forms the coupling between the two basic SOPA amplifiers. The bridge connection allows a doubled output-swing. The major advantage of this configuration however lies in the fact that since both amplifiers are not clocked, the limit cycle oscillations are attracted towards each other. In this way the mean

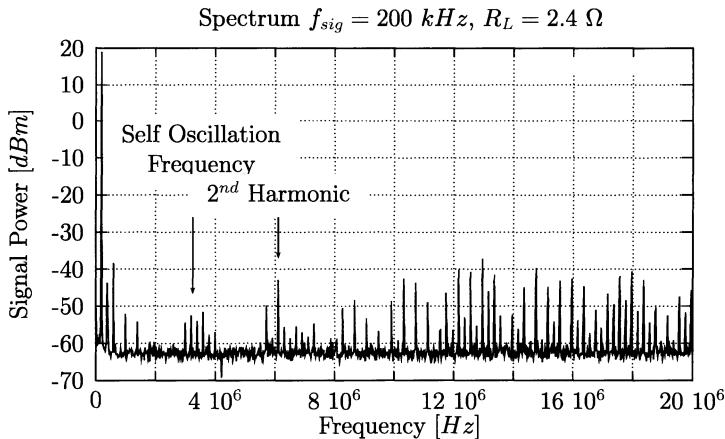


Figure 13.17: Measured output spectrum of zeroth order SOPA line driver with sinusoidal input.

switching frequency becomes common mode for the primary of the line transformer and thus filtered out. Figure 13.17 show a measured output spectrum [21]. The suppression of the limit cycle frequency can be clearly seen. This measurement has been done without extra filtering. Due to this limit cycle attraction technique and the relaxed OSR limits of the modulation, the OSR can be as low as 4.75.

This SOPA line driver has been processed in a .35 μm CMOS technology. The circuit schematic of this implementation is shown in figure 13.18 next to the chip photograph of the actual implementation. The third order loop filter has been constructed as a passive R-C filter. The resistors are thin metal lines, the capacitors are laid out as metal-metal wafers. This allows to generate a very linear filter and makes it possible to process the chip in a standard digital CMOS technology. The digital buffer used is an up-scaled inverter chain in order to minimize the transit time when driving the large output driver. If the delay through the digital buffer becomes larger than the time constant of the loop filter, the linearizing effect of the limit cycle oscillation disappears. The digital buffer is combined with a no-DC-current circuitry by adding the feedback to the NOR-gates. In this way both output transistors are never turned on at the same time, avoiding a direct current from V_{DD} to ground.

Table 13.2 shows the most important measurement results. The SOPA proves to be able to drive the ADSL-Lite flavor with a superior power efficiency.

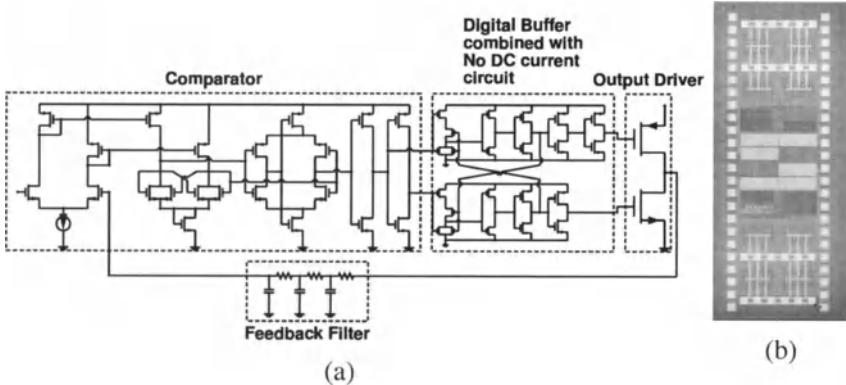


Figure 13.18: Circuit schematic (a) of the implemented SOPA line driver [21] and chip photograph (b).

Table 13.2: Performance Summary of the measured 0th order SOPA.

Parameter	Measured	ADSL-Lite
Technology	0.35µm CMOS	
Supply Voltage	3.3 V	
Output Power	18.7 dBm	16.3 dBm
Voltage Gain	0.9	
-3 dB Bandwidth	800 kHz	500 kHz
Mean switching frequency	3.8 MHz	
Maximum efficiency for sinusoid	61%	
SFDR @ $f_{sig} = 200$ kHz	56.4 dB	
MTPR	41 dB	34 dB

13.4.2.2 Higher order SOPA's

The zeroth order SOPA described in previous section combines a good linearity with high efficiency. The in-band linearity however does not suffice for full ADSL or VDSL systems. In order to reach higher linearities the concept of higher order SOPA's was introduced in [23]. By inserting m integrators in the forward path as illustrated in figure 13.19 for $m = 3$, the characteristics of the modulation are more linearized.

The limit cycle oscillations amplitude and frequency can be recalculated for a m^{th} order SOPA with m integrators with unity gain frequency f_u and a loop filter of order n with cut-off frequency f_c :

$$\omega_{LC} = -\tan\left(\frac{m\pi}{2n}\right)f_c \quad (13.19)$$

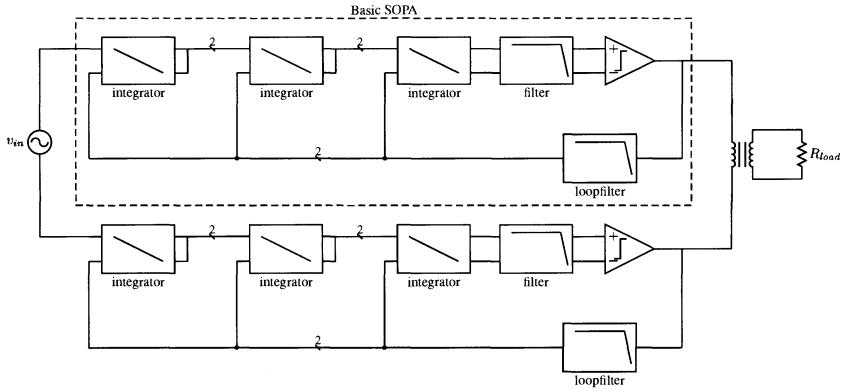


Figure 13.19: Principle schematic of the 3rd order SOPA line driver.

$$A_{LC} = 2^{(1-\frac{n}{2})} \frac{V_{DD}}{\pi} \left(\frac{f_u}{f_c} \right)^m \sin^m \left(\frac{m\pi}{n} \right) \left(1 + \cos \left(\frac{m\pi}{n} \right) \right)^{\frac{n}{2}+m} \quad (13.20)$$

Using these results and the dual-input describing function (13.17) in the closed loop transfer function, the Signal-to-Distortion-Ratio (SDR) can be calculated. Since m and n can only take integer values, a closed formula is too lengthy to give here. Evaluating this formula for $m = 3$ and $n = 4$ gives the SDR (13.21) in function of the cut-off frequencies for the filters and integrators.

$$SDR = 7.5 \cdot 10^{-6} \left(\frac{f_u}{f_c} \right)^3 \frac{V_{DD}^2}{V_{in}} w^3 \left| 1 + \frac{328 i f_c^6}{w^3 (i w + f_c)^3} \right| \times \left| 1 + \frac{328 i f_c^7}{w^3 (i w + f_c)^4} \right| \quad (13.21)$$

In figure 13.20 the SDR is plotted versus frequency for different values of m . The insertion of the integrators will shape the distortion noise. The resulting output spectrum for a sinusoidal input is given in figure 13.21. The noise-shaping can be clearly observed. Thanks to the effect of the oscillator attraction as it has been described in the previous section, more than 40 dB suppression of the limit cycle oscillation has been obtained. Therefor no extra filtering is necessary as compared with a $\Delta\Sigma$ modulated class D amplifier. Only one order filtering is necessary. The transfer characteristic of the line driver itself contains a second order characteristic, so maximal efficiency can be obtained.

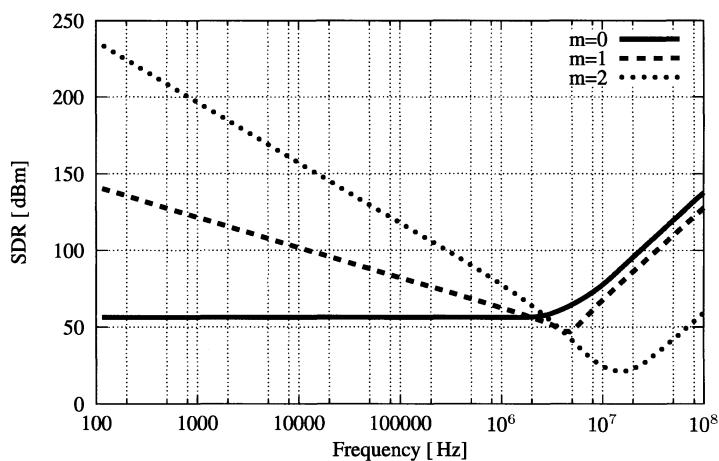


Figure 13.20: SDR versus frequency for different number of integrators m .

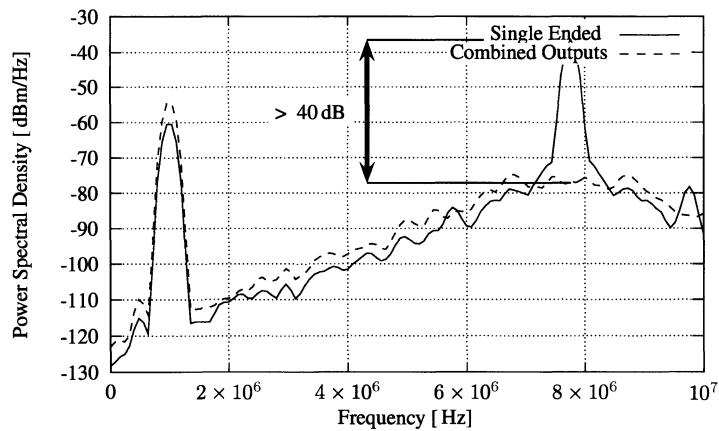


Figure 13.21: Output spectrum for a sinusoidal input of a 3rd order SOPA.

Since the bandwidth is governed by the the limit cycle frequency ω_{LC} and this frequency is controlled by the loop filters cut-off frequency f_c (13.19) and the linearity is governed by the integrators, the design problem is decoupled.

For the circuit level design the following has to be taken into account :

- Like in the zeroth order case, the transit time through the comparator and digital buffer should be minimized. Since the comparator is not clocked, it cannot be placed in a a-stable position. A fast continuous time comparator needs to be designed. Positive feedback can be utilized to fasten the comparison. When constructing a line driver for VDSL applications, this can cause a technological limit.
- The linearity and noise requirements for the integrators are relaxed from the systems inputs towards the comparator inputs. This is due to the noise-shaping of the integrators non-linearity by the preceding integrators in the loop. The linearity of the first integrator however will govern the overall line drivers distortion. Therefor in the design of [23] the integrator is linearized by resistive source degeneration (fig. 13.22). The generated distortion of the first integrator can then be calculated as :

$$HD_3 = \frac{-1}{32} \frac{g_m^2}{I_{Bias}^2 (g_m R_E + 1)^3} \quad (13.22)$$

The power consumption of the integrators will lower the maximal obtainable efficiency, so matching characteristics of the used technology should be taken into account to minimize the power consumption without losing performance.

- The linearity and noise generation of the loop filter will directly appear at the output. A simple RC filter however will suffice.

A 3rd order SOPA has been processed in a .35 μm CMOS technology [23]. A chip photograph is given in figure 13.23. The large output drivers can be clearly seen, since they take more than a third of the chip area. Enough care has to be taken to avoid electro-migration of the metal. The bottom half of the chips is taken by the integrators. The up-scaling of the integrators from input towards the comparator is noticeable by the increase in capacitor area from the bottom of the chip towards the center. To protect the sensitive analog parts from supply noise, decoupling capacitances have been placed as close as possible to every analog part. The chip has been measured and its performance as a central office ADSL and VDSL modem has been verified. For the ADSL verification a Missing Tone Power Ratio test has been performed by applying a downstream ADSL signal to the inputs of the line driver. to measure the MTPR some tones are left out to act as antenna-tones. In

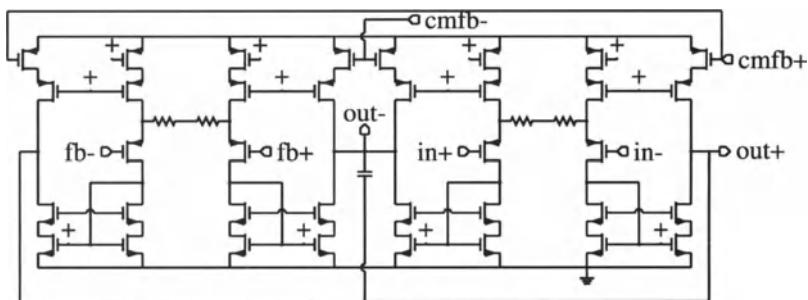


Figure 13.22: Integrator used in [23]. Resistive source degeneration is utilized to meet linearity specifications.

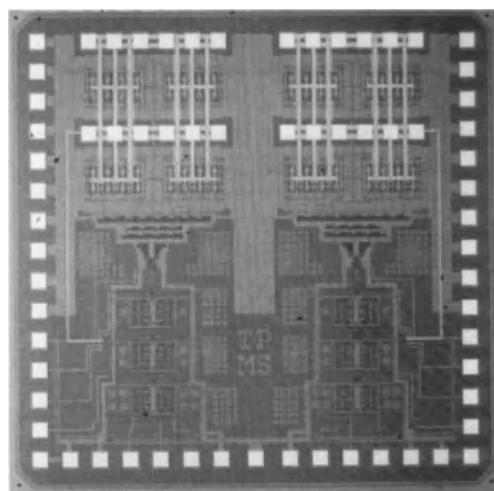


Figure 13.23: Chip photograph of the SOPA line driver presented in [23].

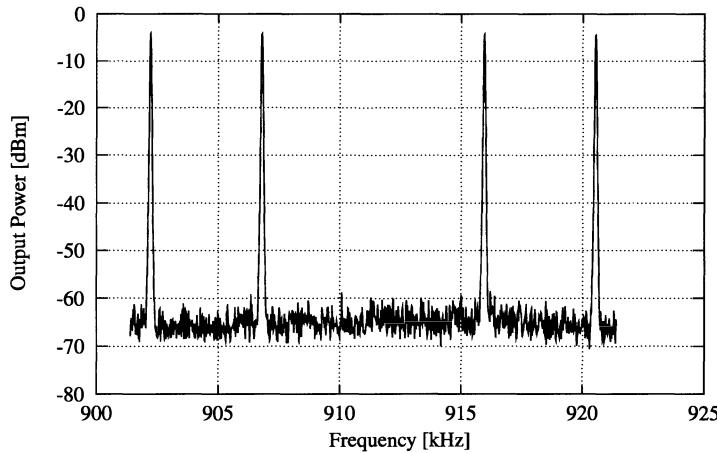


Figure 13.24: MTPR measurement of the SOPA presented in [23] around the 226th tone.

figure 13.24 the measurement around a high frequent antenna tone has been given since these tones will provide the worst case measurement.

Table 13.3: Performance summary of the measured 3rd order SOPA.

Parameter	Measured	xDSL specs
Technology	0.35 μ m CMOS	
Supply Voltage	3.3 V	
Bandwidth	> 8.6 MHz	8.5 MHz (VDSL DS)
MTPR	56 dB	55 dB
Output Power ADSL	-38.7 dBm/Hz	-40 dBm/Hz (DS)
Out of Band PSD	<-103 dBm/Hz	-100 dBm/Hz
Crest Factor	> 5	>5
Efficiency for 100 mW ADSL	47 %	

An MTPR of 56 dB has been achieved for a 100 mW ADSL test signal with an efficiency of 47 %. For a signal of 134 mW a 53 % power efficiency has been measured. The line driver is also able to send VDSL upstream signals with a bandwidth of 8.5 MHz. Table 13.3 summarizes the main measured specifications of this SOPA line driver.

By going to higher order SOPA line drivers, the opportunity is opened to design highly efficient line drivers that are able to driver signals with a high bandwidth and a high crest factor. Since the power consumption of the analog integrators will

degrade the power efficiency, much care has to be taken to select the minimal order SOPA that suffice to meet the required specifications.

13.5 CONCLUSION

The emerge of the Internet and the resulting fast developments in the area of wire-line communications has put the line driver design back in full spotlights. The demanding specifications of xDSL systems for the power amplifiers in terms of bandwidth, linearity and efficiency has put the research in this field in a higher gear. Class AB line drivers are well known and thoroughly studied power amplifiers that are mostly used for their high linearity. Recent research in the area of the quiescent current control of these type of power amplifiers has moved the distortion specifications towards the ones of the class A operating region, while pushing the power efficiency towards the class B operating region. For modern xDSL flavors that use the discrete multi-tone modulation the crest factor is above 15 dB. The class AB output stage has for this applications thus the disadvantage that its efficiency is inverse proportional with the crest factor. Even for ideal class B operation, which is the lower limit for power dissipation, this will lead to a maximum efficiency of 22 % for an ADSL signal, which is unacceptable for a modem at the central office. Since this efficiency problem is related with the ratio between the rms voltage of the signal and the supply voltage, the class G solution emerges. By signal-dependent switching between several supplies a power amplifier can be constructed that has the possibility to have class AB linearity with more than double the efficiency. This approach has proven to be very successful in the audio amplifier field but the high bandwidth requirements of an xDSL signal make class G a real challenge. Up to now research has proven to make class G line drivers that are comparable with state-of-the-art class AB line drivers in terms of linearity, bandwidth and efficiency, but the class AB specs were not beaten yet. It was shown in this chapter that the further elaboration of the modulated supply technique, being the class H power amplifier, will for the present state of the technology not be a beneficial choice.

Switching type line drivers have the theoretical possibility to have a 100 % efficiency, independent on the crest factor, since an ideal switch does not dissipate power. The clocking frequency is the limiting factor. Due to the high bandwidth of current wire-line applications and the high over-switching ratios needed for highly linear PWM signals, the power switches need to be switched at too high frequencies. On the other hand, a high OSR will facilitate the design of the steep output filter. This filter will become bulky and should be avoided. Much research is carried out to improve PWM modulation schemes to meet the high linearity specifications. In order to reach a high efficiency, high linearity line driver for wire-line communications, novel architecture should be researched. A Self Oscillating Power

Amplifier has been described as a possible new topology. By using a asynchronous modulation scheme a highly linear switching signal is obtained. The problems related with output filtering are relaxed by more than 40 dB due to the effect of oscillator attraction.

REFERENCES

- [1] Damien Macq, "DMT ADSL Circuits and Systems", in *ISSCC : xDSL Broadband Interactive Communications*, February 1998.
- [2] Bill Schweber, "Analog Front Ends Bridge the xDSL-to-real-world chasm", *EDN magazine*, pp. 48–64, April 1999.
- [3] Kim Maxwell, "Asymmetric Digital Subscriber Line : Interim Technology for the Next Forty Years", *IEEE Commun. Mag.*, pp. 100–106, October 1996.
- [4] Leon Cloetens, "Broadband Access : The Last Mile", in *ISSCC Digest of Technical Papers*, pp. 18–21, February 2001.
- [5] Russell Apfel, "ADSL Transceivers : Optimizing Power, Cost, Function and Performance", in *ISSCC : SSCTC Workshop on Analog Telecom Access Circuits and Concepts*, February 2002.
- [6] Roger Benton, Russell Apfel, Bruce Webb, Jerome Wenske, Walt Schopfer and Frank Thiel, "A High-Voltage Line Driver (HVLDR) for Combined Voice and Data Services", in *ISSCC Digest of Technical Papers*, pp. 302–303, February 2001.
- [7] Domenico Rossi, "Silicon Technologies and Circuit Topologies for High Efficiency in ADSL Line Drivers", in *ISSCC : SSCTC Workshop on Analog Telecom Access Circuits and Concepts*, February 2002.
- [8] Irving Kalet, "The Multitone Channel", *IEEE Trans. Commun.*, vol. 37, n. 2, pp. 119–124, February 1989.
- [9] D.J.G. MestDagh, P.M.P. Spruyt and B. Biran, "Effect of Amplitude Clipping in DMT-ADSL Transceivers", *Electronics Letters*, vol. 29, n. 15, pp. 1354–1355, July 1993.
- [10] Jose Tellado-Mourelo, *Peak to Average Power Reduction for Multicarrier Modulation*, PhD thesis, Department of Electrical Engineering, Stanford University, September 1999.
- [11] M. Cresi, R. Hester, K. Maclean, M. Agah, J. Quarfoot, C. Kozak, N. Gibson and T. Hagen, "An ADSL Central Office Analog Front-End Integrating Actively-Terminated Line Driver, Receiver and Filters", in *ISSCC Digest of Technical Papers*, pp. 304–305, February 2001.
- [12] Mark Ingels, Sena Bojja and Patrick Wouters, "A 0.5 μm CMOS Low-Distortion Low-Power Line Driver with Embedded Digital Adaptive Bias Algorithm for Integrated ADSL Analog Front-Ends", in *ISSCC Digest of Technical Papers*, pp. 324–325, February 2002.

- [13] Faramarz Sabouti and Reza Shariabdoust, “A 740 mW ADSL Line Driver for Central Office with 75 dB MTPR”, in *ISSCC Digest of Technical Papers*, pp. 322–323, February 2002.
- [14] H. Casier, P. Wouters and D. Sallaerts, “A 3.3-V Low-Distortion ISDN Line Driver with a Novel Quiescent Current Control Circuit”, *IEEE J. Solid-State Circuits*, vol. 33, n. 7, pp. 1130–1133, July 1998.
- [15] Haideh Khorramabadi, “A CMOS line driver with 80 dB linearity for ISDN applications”, *IEEE J. Solid-State Circuits*, vol. 27, n. 4, pp. 539–544, April 1992.
- [16] John Pierdomenico, Scott Wurcer and Bob Day, “A 744 mW Adaptive Supply Full-Rate ADSL CO Driver”, in *ISSCC Digest of Technical Papers*, pp. 320–321, February 2002.
- [17] Marcus W. May, Michael R. May and John E. Willis, “A Synchronous Dual-Input Switching dc-dc Converter Using Multibit Noise-Shaped Switch Control”, in *ISSCC Digest of Technical Papers*, pp. 358–359, February 2001.
- [18] R. A. R. van der Zee and A. J. M. van Tuijl, “A Power Efficient Audio Amplifier Combining Switching and Linear Techniques”, in *Proceedings of the 24th European Solid-State Circuits Conference*, pp. 288–291, 1998.
- [19] Brian E. Atwood, “Design parameters important for the optimization of very-high-fidelity PWM (class D) audio amplifiers”, *Journal of the audio engineering society*, vol. 31, n. 11, pp. 842–853, November 1983.
- [20] Enrico Dallago, “Advances in High-Frequency Power Conversion by Delta-Sigma Modulation”, *IEEE Trans. Circuits Syst. I*, vol. 44, pp. 712–721, August 1997.
- [21] Tim Piessens and Michiel Steyaert, “SOPA : A High-Efficiency Line Driver in $0.35\mu\text{m}$ CMOS using a Self-Oscillating Power Amplifier”, in *ISSCC Digest of Technical Papers*, pp. 306–307, February 2001.
- [22] Arthur Gelb and Wallace Vander Velde, *Multiple-input Describing Functions and Nonlinear System Design*, McGraw-Hill Book Company, 1968.
- [23] Tim Piessens and Michiel Steyaert, “A Central Officed Combined ADSL-VDSL Line Driver in $.35\mu\text{m}$ CMOS”, in *Proceedings Custom Integrated Circuits Conference*, pp. 45–48. IEEE, May 2002.

Chapter 14

PGAs and Filters

Manuel Delgado-Restituto and Angel Rodríguez-Vázquez

14.1 INTRODUCTION ^{†1}

Amplifiers and filters are commonplace devices in the *analog front-end* (AFE) of communication transmitters-receivers (transceivers, in short). In a general sense, these devices provide the necessary adaptation, in terms of power adjustment and signal isolation, between the transmission media (e.g., atmosphere, free space, cable, twisted-pair, optic fiber) and the digital signal processor (DSP) which performs most of the algorithmic tasks needed to guarantee a reliable transmission/reception of the information [1]. In some cases as, for instance, in wireless transceivers, amplification and filtering may take place at multiple steps along the AFE; often using different technologies (CMOS, silicon bipolar, GaAs) or external passive components (e.g., surface-acoustic wave filters) depending on the frequency range at which operations are realized [2]. In this chapter, following the main stream of the book, we focus on the realization of those amplifiers and filters which are used to drive signals to/from the AFE data converters^{†2} at the interface with the DSP block, paying special attention on their implementation in inexpensive CMOS technologies. Such amplifiers and filters are symbolically shown in Fig. 14.1, where preceding/following circuits for reception/transmission have been globally called Rx/Tx medium interface, respectively.

Signal conditioning done by baseband amplifiers and filters has different objec-

-
1. **Acknowledgement:** This work has been partially funded by the spanish MCyT under Projects TIC2001-0929 (ADAVERE) and FIT-070000-2001-843 (AFIN).
 2. In wireless applications, these amplifiers and filters, together with data converters themselves, constitute the *analog baseband* section of the transceiver. The term baseband emphasizes the fact that no frequency shifting of the information-bearing signal transferred to or delivered from the DSP block is accomplished at this point. Though, it can not be rigorously extrapolated to other communication scenarios where no frequency translation takes place, e.g., wireline, we will refer to the analog amplifiers and filters besides the transceiver data converters as baseband.

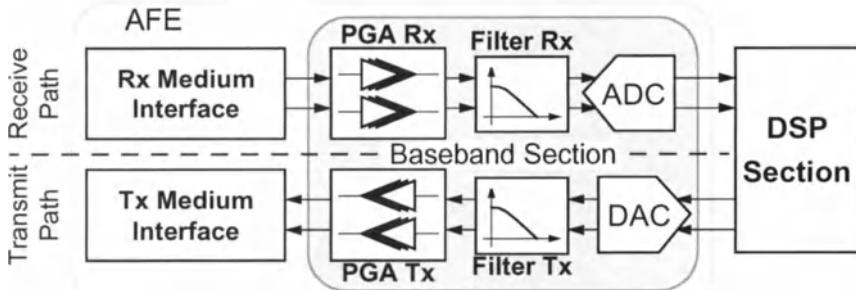


Figure 14.1: General transceiver architecture showing the amplifiers and filters covered in this chapter. Double arrows symbolize the more general case of in-phase and quadrature (I/Q) signal processing.

tives depending on whether they are used in the receiver or the transmitter paths of the AFE. In the receiver path (elements above the dotted line in Fig. 14.1), the main task is to attenuate all the unwanted signals (generally called interferers^{†3}) outside the desired information channel, while adjusting the output signal amplitude to a safe level below the full-scale range of the following analog-to-digital converter (ADC). Given the presumably changing propagation conditions, signal level adjustment must be realized by an electrically controlled programmable gain amplifier (PGA). Most often, the gain setting of the PGA is adaptively updated from the DSP block. Also, because of the sampled-data nature of the ADC, receiver filters must conveniently attenuate any energy that could potentially alias into the desired channel as a result of sampling. For similar reasons, baseband filters and amplifiers must be linear enough so that out-of-channel interferers do not create intermodulation distortion signals which fall in the desired channel.

In the transmitter side (elements below the dotted line in Fig. 14.1), the main objective is to smooth the output of the preceding digital-to-analog converter (DAC), paying special attention on lowering the amount of distortion incurred on adjacent channels in order to not degrade appreciably the link performance. In addition, filters and amplifiers must be used to attenuate the out-of-channel noise in order to reduce the filtering requirements at later stages of the transmitter.

-
3. In a wireless scenario, interferers are all those unintended signal transmissions which leak energy in the frequency band allocated for the standard and which can be produced by other users sharing the application, electromagnetic appliances (a well-known example is the radiation emitted by microwave ovens on the 2GHz band) or even natural phenomena. In wireline applications as, for instance, in asymmetric digital subscriber line (ADSL) transmission units, crosstalk between upstream and downstream transmitted signals due to electromagnetic coupling in the line is the major source of interference.

Apart from the application itself –commonly standardized by regulatory bodies such as ITU, ANSI, or ETSI, among others– design specifications for the amplifiers and filters in order to achieve the aforementioned objectives strongly depend on the transceiver topology in which they are embedded. In fact, design aspects, which are crucial in certain architectures, have almost no impact in others, and this obviously affects the proposal of circuit solutions. Thus, before dealing with circuit techniques for the physical realization of baseband filters and amplifiers, it is necessary to explore their design requirements from a transceiver perspective and show how such specifications can be extracted from both the communication standard and the particular transceiver topology selected for integration. This is, indeed, the main subject of this chapter.

The chapter will cover the first two steps of a typical top-down design procedure, namely, the transmission of specifications from the system (transceiver) level to the overall baseband section, and from here to its building blocks. Last steps of the hierarchical design approach, concerning the circuit and physical levels, are described in excellent monographs and reprint volumes such as [3]-[6]. Also, the more algorithmic aspects for filter design (approximation techniques, passive-to-active circuit transformations, node scaling, and so on) are described in classical textbooks, such as [7], [8].

Section 14.2 analyzes the different trade-offs emerging in the first step of the top-down approach and illustrates with exemplary wireless receiver prototypes reported in the literature the wide variety of high-level requirements for baseband sections that can be found in practice. Then, Section 14.3 reviews different aspects for the block-level realization of the baseband section. Namely, it surveys the alternatives that can be found with regard to the channel select filtering scheme used, the type of signals which are handled, the number of stages required to totally or partially isolate/amplify the desired information channel, and the granularity of the gain adjustment accomplished by the PGA.

Finally, Section 14.4 gives some concluding remarks, highlights the current trends on the implementation of PGAs and filters and provides the readers with up-to-date references for further exploring most recent developments on the topic.

14.2 PGAS AND FILTERS FROM A TRANSCEIVER PERSPECTIVE

Any electronic transceiver, and hence, the PGAs and filters included therein, must be designed so as to exactly comply with the specifications of the targeted application standard. However, standards do not give explicit recommendations for the physical realization of the transceiver but, rather, consider it as black box, and outline a set of evaluation tests from which overall performance parameters of the transceiver must be inferred.

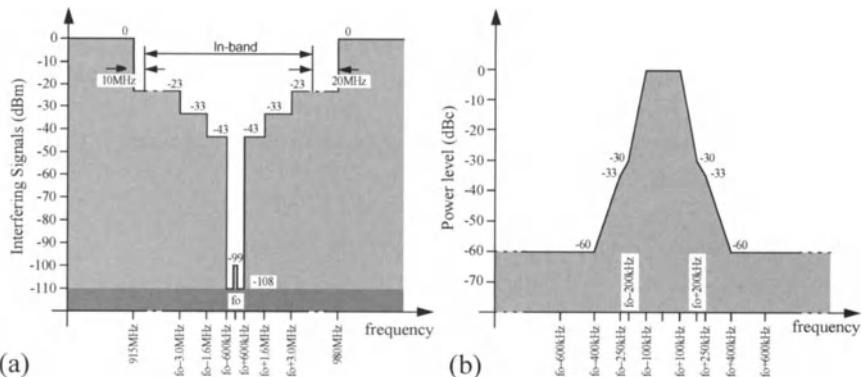


Figure 14.2: (a) GSM blocking profile. GSM also specifies maximum interfering signals for the adjacent (at $f_0 \pm 200\text{kHz}$) and alternate (at $f_0 \pm 400\text{kHz}$) channels of 9 and 41dB above the wanted signal level, respectively. (b) Close-in GSM output spectral mask for a maximum peak output power of +33dBm (measurement resolution bandwidth must be 30kHz).

Tests to evaluate the reception and transmission capabilities of the transceiver are essentially different. Receiver tests describe different propagation conditions in the communication channel, that depend on the particular performance which is evaluated, and assesses if the receiver meet a minimum reliability performance, usually expressed in terms of a maximum error rate value^{t4}. In this way, standards propose specific tests to validate receiver aspects such as *sensitivity* (minimal signal at the input which guarantees a prescribed carrier-to-noise-and-interference ratio, $C/(I+N)_{min}$, at the output^{t5}), *selectivity* (ability to separate the desired channel from unwanted signals received at other frequencies – maximum power levels of interferers as a function of the frequency offset from the desired channel define the *interference and blocking profile* of the receiver-) and *linearity* (tolerance to nonlinear effects, such as intermodulation distortion) [9], [10]. As an example, Fig. 14.2(a) shows the blocking profile for the GSM standard [14], and

-
- 4. Most often, the error rate measure used to quantify the reliability of the transmission link is the bit error rate (*BER*), however, other related measures can be also specified. An example is the maximum clipping rate imposed by the asymmetric digital subscriber line (ADSL) standard [11].
 - 5. To satisfy a given error rate, the carrier-to-noise-and-interference ratio at the input of the demodulator (output of the AFE receiver path) must be larger than a minimum value $C/(I+N)_{min}$, which depends on the particular modulation scheme proposed by the standard. For instance, in the global system mobile (GSM) standard, which uses GMSK modulation, $C/(I+N)_{min} = 9\text{dB}$, and in the ADSL standard, which employs DMT modulation, $C/(I+N)_{min}$ rises to 55dB [11]-[13].

Table 14.1: Summary of GSM receiver specifications

Performance Metric	Spec
Reference sensitivity (dBm)	-102
Input noise floor, N_{floor} (dBm)	-111
Available thermal noise, N_r (dBm)	-120.85
Noise figure, NF (dB)	9.85
Input range, iR (dB) ^a	87
Input-referred IP3, iIP_3 (dBm)	-18
1-dB Compression point, iCP (dBm)	+5
Spurious-free dynamic range, $SFDR$ (dB)	+53

a. Input range is here defined as the maximum desired signal power (-15dBm in GSM) that the receiver must be able to detect, referred to the sensitivity level.

Table 14.1 summarizes different (input-referred) receiver requirements derived from the sensitivity and linearity tests. Definitions and procedures leading to Table 14.1 are detailed in Appendices 14.I, 14.II and 14.III which focus, respectively, on the performance metrics related to sensitivity, linearity and dynamic range.

With regard to the transmitter, performance tests define different operation conditions (single tone or modulated baseband signal transmissions, switching transients during power up and down of the transmitter) and check if the output signal verifies specific constraints regarding *modulation accuracy* and *power spectrum purity*. Estimation of the modulation accuracy is typically performed using a modulated baseband signal and measuring the deviations in phase and magnitude between the ideal response and the actual transmitted signal. Regarding spectral purity, corresponding tests are passed whenever the power emission of the transmitter, including the wanted modulated signal, as well as, the transmitted noise and any other spurious response due to circuit nonidealities, fall below an *output spectral mask* defined by the standard. As in the receiver, the above restrictions also translate into minimum requirements for the sensitivity, selectivity and linearity of the circuits in the transmitter chain. For illustration purposes, Fig. 14.2(b) shows the output spectral mask specified by the GSM standard.

From the above global requirements (referred to the communication channel), designers are then faced to, first, select a system architecture able to fulfill all the reception/transmission tests, and, second, identify design specifications for each

of the blocks building up the chosen circuit topology. These aspects are, respectively, discussed in the following two subsections.

14.2.1 System architectures

Though there exist classical transceiver architectures whose efficiency and adaptivity to different specifications have been largely contrasted during years and that, even today, can be found in the vast majority of commercial communication appliances, is not unusual that new variants, or even early discarded architectures, are (re-)introduced in the literature which, fostered by most recent technological advances, can be potentially suited for an overall cost reduction of the product or a significant system performance improvement.

An example of this scenario (and the resulting architecture diversity) is found in the implementation of wireless receivers [15]-[24] (a similar situation occurs with radio transmitters [22], [26]-[27]). Pushed by the system-on-a-chip (SoC) trend in microelectronics, different proposals have been recently made which aim to provide a higher degree of integration than traditional superheterodyne radio receivers [15]-[17]. In superheterodyne schemes –see Fig. 14.3(a)–, image^{†6} and large blocking signals rejection are provided by high Q, high performance off-chip filters, what leads to high size and power consumptions (interface to off-chip components requires to transforming down the local on-chip impedances, usually to 50Ω). As a counterpart, specifications for the baseband section in these receivers are largely relaxed and low to moderate dynamic range performance blocks are enough to digitize the selected channel. The first column of Table 14.2 illustrates this point by indicating the system requirements at the input of the baseband section, in a recently published superheterodyne CMOS GSM receiver [17].

Alternatively, single chip solutions (two examples are shown in Figs.14.3(a) [23] and (b) [24]) eliminate both the image rejection (IR) and intermediate frequency (IF) filters, thus, avoiding the need to take sensitive nodes out of the chip (influence of package parasitics are minimized) and reducing the size and cost of the overall receiver [18]-[24]. The price of this increased integrability is that functions originally realized by the suppressed discrete components must be now implemented on-chip. Proposed solutions for the image rejection problem relies on signal cancellation techniques based on in-phase (I) and quadrature (Q) signal

-
- 6. The problem of image signal is inherent to the operation of frequency translation realized by mixers. Down-conversion mixers operate on the principle that if two sinusoidal signals are multiplied together, the resulting product has sum and difference frequency components. As a consequence, if the oscillator frequency is set to f_{of} , there are two input frequencies, $f_{of} \pm f_{if}$, that are translated to the *same* intermediate frequency, f_{if} . Typically, one of these frequencies is the desired signal while the other is commonly called the image signal.

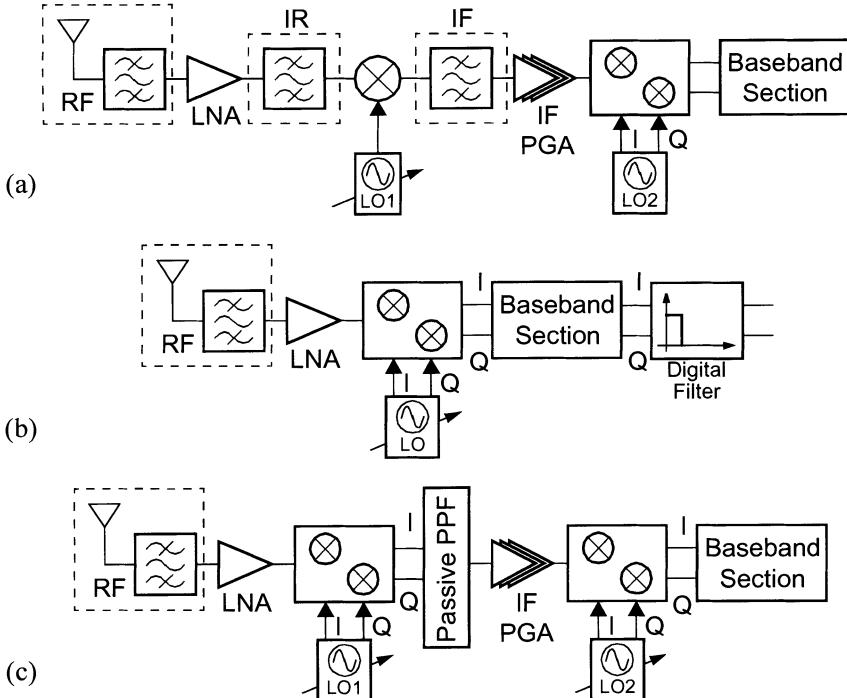


Figure 14.3: Wireless receiver architectures. (a) Superheterodyne dual-downconversion; (b) Low-IF (or zero-IF) single-downconversion; (c) Wideband IF low-IF dual-downconversion. Blocks surrounded by dashed lines are off-chip. In all three cases, the RF filter (most often, a duplexer) is a passive element designed to pass energy in the receive frequency band (935 to 960 MHz for GSM), while attenuating energy outside.

representations and the use of sophisticated blocks such as quadrature double-balanced mixers or complex polyphase filters [22]. The precision with which the I and Q signal paths can be matched determines how good the mirrored signal can be rejected. On the other hand, as the entire in-band spectrum is translated to baseband, the problem of blocking signals suppression must be tackled by putting more severe requirements on the selectivity, linearity and dynamic range of the baseband circuits. This is clearly observed by comparing the first column of Table 14.2 with the remaining second and third columns, the second one corresponding to a recently reported low-IF –Fig. 14.3(b)– DCS-1800 receiver^{†7} [23]; and the third one related to a double conversion low-IF receiver for GSM [24]. It should be noted that this performance level increment is partly due to effects, almost neg-

Table 14.2: Input-referred requirements for different wireless receiver baseband sections.

Signal level/Performance metric	SH GSM [17]	Low-IF GSM [23]	Dual Conv. Low-IF GSM [24]
Noise figure, NF (dB)	NA	7.8	26
Intercept Point, $iIP3$ (dBm)	25	27.4	25
Compression Point, iCP (dBm)	10	28	NA
Filter Order - Approximation Technique	4 Bessel	1	5 Tchebyshev
Cutoff Frequency (kHz)	150	300	20-260
Gain@Steps (dB)	16	19-43@6	-2-58@1.875
Output capability (V _{pp})	1	2	1.2

ligible in superheterodyne schemes, but increasingly relevant in single-chip solutions. Examples, specially acute in zero-IF receivers, are signal-dependent DC offset resulting from LO leakage at the mixer output, second order intermodulation and flicker noise [9], [18].

If, as already shown, implementation scenarios may considerably vary even when targeting the same standard, the requirements and circuit solutions for baseband sections explode when dealing with different applications. Obviously, design difficulties, in particular, using CMOS technologies, are tied to the required system specifications, and, for this reason, the availability of CMOS demonstrators has been gradually demonstrated in accordance to the projected performance. First demonstrators focused on low-performance wireless paging applications (such as ERMES), followed by short-range communicators for cordless (such as DECT), positioning (such as GPS) or wireless local area network (such as Bluetooth) applications. Single-chip full-CMOS realizations for the much more challenging narrow-band cellular applications, such as GSM, have been only very recently reported in the literature [23], [24].

Achievement of high performance baseband sections for CMOS single-chip transceivers, in particular meeting very demanding communication standards, out-

7. The digital cellular system DCS-1800 is a frequency-shifted version of the GSM standard, with slightly more relaxed radio interface specifications.

lines the search of innovative circuit solutions and the efficient partitioning of tasks among building circuits. These aspects will be covered in Section 14.3.

14.2.2 Functional blocks characterization

After architecture selection, the functional blocks of the transceiver must be individually characterized. The procedure to achieve this goal can be formulated as an optimization problem in which variables are the design parameters of the functional blocks; constraints are imposed by the input-referred specifications defined by the standard, the different issues affecting system performance (image frequencies, distortion, intermodulation, aliasing, etc.) or any other extra limitation such as the availability of discrete components; and objectives are, most often, the minimization of the area and power consumptions of the solution [22]. The overall set of design specifications for the building blocks define the so-called *system planning*. Data in Table 14.2 actually correspond to partial views, focused in the baseband section, of different wireless receiver plannings.

Design parameters of the functional blocks can be classified into general and specific. General parameters must be defined for every block irrespective of its intended operation. They include [22]:

- Operating frequency.
- Bandwidth over which wanted operations must be realized.
- Amplification within the passband of the block. This can be expressed by the voltage gain A_j , (available) power gain G_{aj} or any other similar metric, upon convenience.
- Noise figure, NF_j .
- Nonlinearity, commonly expressed by the input-referred second- and third-order intercept points, $iIP2_j$ and $iIP3_j$, and the input-referred 1-dB compression point iCP_j (see Appendix 14.II for details).

On the other hand, specific parameters are those associated to the particular functionality of the block. For instance, specific design parameters for PGAs are the variation range and sweeping steps of the programmable gain and, in the case of filters, specific parameters are those related to its transfer characteristics, essentially, cut-off frequencies, passband ripple and stopband attenuation.

As already mentioned, any valid system planning must conform with the power levels for the wanted channel, noise and interfering signals specified by the standard. Actually, the system planning is usually accompanied by a level diagram which shows how the different signals evolve along the receiver/transmitter chain. From this diagram, the dynamic range specifications DR_j for each of the building blocks can be easily deduced. As an example, Fig. 14.4 shows the level diagram associated to the GSM receiver of the first column in Table 14.2.

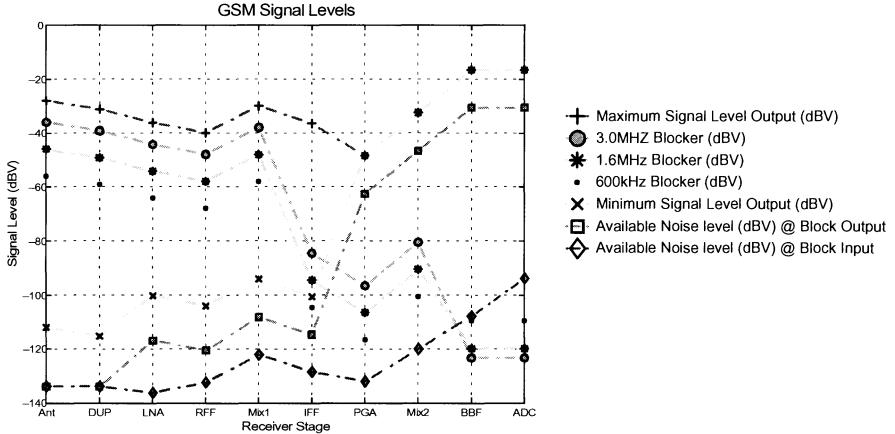


Figure 14.4: Level diagram for the wireless GSM receiver in [17].

A fundamental aspect on the development of the system planning is that specifications of a single block can not be changed independently, without altering the required performance for other elements. Moreover, such mutual influence is not uniform along the chain but, depending on the particular performance metric, building block specifications become more and more critical as we move up or down the system cascade. This can be easily shown by observing how the noise figures and intercept points of individual blocks determine same (input-referred) parameters of the overall receiver/transmitter. For a cascade circuit with m elements, the overall noise factor F_{tot} (noise figure expressed in ratio) can be calculated from Friis' formula as (see Appendix 14.I)

$$F_{tot} = 1 + (F_1 - 1) + \frac{(F_2 - 1)}{g_{a1}} + \dots + \frac{(F_m - 1)}{g_{a1} \cdots g_{a,m-1}} \quad (14.1)$$

where g_{aj} and F_j are, respectively, the (available) gain and noise factor of the j -th block. Equation (14.1) shows that the gain and noise figure of the first stage are critical in achieving a low overall noise figure, because the contribution of each subsequent element is reduced by the total available gain power preceding it.

Similarly, it can be shown that the linearity (represented by $iIP3$) of a cascade of signal blocks can be approximated as (see Appendix 14.II),

$$\frac{1}{iip_{3,tot}} \approx \frac{1}{iip_{3,1}} + \frac{g_1}{iip_{3,2}} + \frac{g_1 g_2}{iip_{3,3}} + \dots + \frac{g_1 g_2 \cdots g_{m-1}}{iip_{3,m}} \quad (14.2)$$

where g_j and $iip_{3,j}$ are, respectively, the power gain and input 3rd order intercept point of the j -th block and, it is implicitly assumed that the distortion contributions from each of the blocks are uncorrelated. By re-expressing (14.2) in terms of output referred intercept points, we can demonstrate a certain symmetry with (14.1),

$$\frac{1}{oip_{3,tot}} \approx \frac{1}{oip_{3,m}} + \frac{1}{g_m oip_{3,m-1}} + \frac{1}{g_m g_{m-1} oip_{3,m-2}} + \dots \quad (14.3)$$

where it can be observed that the contribution of each stage to the total $oip_{3,tot}$ is reduced by the gain that follows. Thus, the last circuit element in the chain tends to contribute the most to the distortion. As will be shown in Section 14.3, both (14.1) and (14.3) will play an important role on the design of optimum baseband sections.

Another essential aspect on the search of an optimum system planning is the formulation of cost functions for each of the building blocks, so that their area and power consumptions (or any other variable considered in the optimization objectives) can be estimated in terms of the associated general and specific design parameters. For instance, in the case of active integrated filters, area and power consumptions can be estimated in terms of the required dynamic range, the quality factor and the operation frequency [3]. From these cost functions, which must obviously account physical and technological limitations, as well as, any other design constraint such as temperature or voltage supply conditions, an overall evaluation of the system planning can be obtained.

14.3 BASEBAND SECTION ARCHITECTURES AND TECHNIQUES

Given the wide variety of baseband requirements that can be found in practice – as already mentioned, they depend upon the targeted standard, chosen system architecture and splitting of specifications among functional blocks– circuit solutions for the baseband section are seemingly quiet different. In this section, four design aspects which should be evaluated towards system implementation are reviewed. The first one is related to the circuit strategy (fully-analog or mixed-signal) used for channel isolation. The second aspect relies on the type of signal (real or complex) handled. Another aspect is to define the number of stages required to perform the amplification and filtering functions. Finally, a distinction must be made between those architectures in which amplification programming is realized in discrete steps or in a continuous manner.

14.3.1 Channel Isolation Techniques

By channel isolation, we mean the process by which input signals other than the desired channel (interferers and image frequencies in receivers, or spurious DAC responses in transmitters) are attenuated to levels below the noise floor of subsequent stages –quantization noise, in the case of digital blocks–.

As already mentioned, channel isolation techniques can be classified into fully-analog and mixed-signal. While the filtering requirements are the same for either of the two methods, the choice of fully-analog or mixed-signal selection impacts the dynamic range requirements of the data converter and the programmability of the transceiver. For the sake of conciseness, the following discussion will focus exclusively on receivers, although main conclusions can be also extrapolated to the transmitter side – called references provide examples for both cases.

14.3.1.1 Fully-Analog method

In this case, channel isolation is fully performed prior to the ADC. Therefore, only a low resolution converter (typically 8 bits or less) with enough bandwidth to digitize the desired channel is required. Conversely, the analog filter must have very high dynamic range and linearity to select the desired channel in the presence of strong adjacent channel interferers.

Figs.14.5(a) and (b) show two possible implementations of the fully-analog channel isolation concept. That in Fig. 14.5(a) employs continuous-time (lowpass or bandpass) analog filtering in front of the ADC [28]-[36]. Because the required selectivity is usually high, filters also prevent against anti-alias effects from the ADC sampling process. Depending on the linearity requirement and the operation frequency, such filters can be implemented using transconductance- C (also called, g_m - C) or active RC techniques. g_m - C filters can operate at very high frequencies, but their $iIP3$ is low, essentially limited by transconductor nonlinearity. By contrast, active RC filters exhibit very high linearity, but their use is limited to the low-MHz frequency range.

One key issue in the design of continuous-time filters is the compensation for circuit component variations. Since the filter frequency response depends on the absolute value of capacitors and (trans-) conductances, and they suffer from large statistical variations, filter poles may deviate by more than $\pm 30\%$ from design intent and, therefore, some form of on-chip tuning is usually required. This topic is thoroughly revised in the selected reprint volume [3], where different automatic tuning schemes for both the passband frequency and quality factor of continuous-time filters are proposed.

Another possibility for fully-analog channel isolation, shown in Fig. 14.5(b),

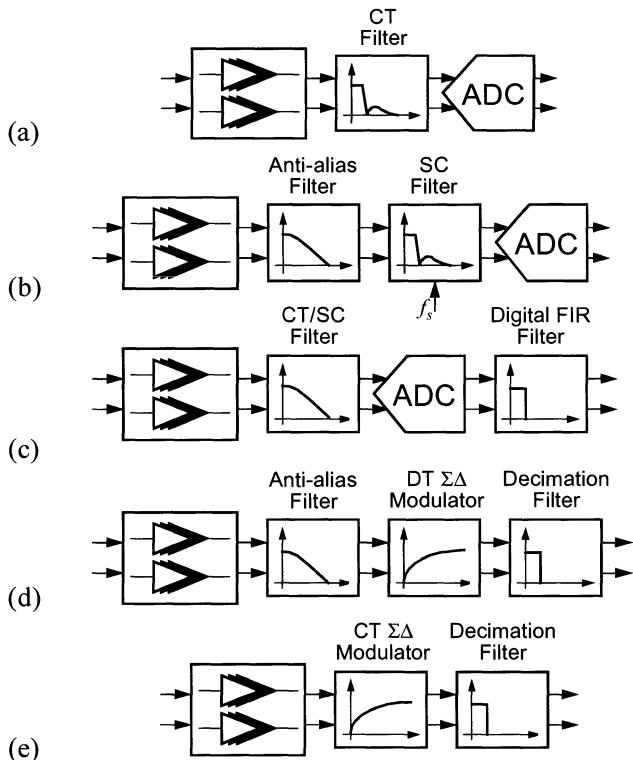


Figure 14.5: Alternative baseband receiver sections. In (a) and (b), the information channel is fully selected by the analog baseband strip. They differ in the applied filtering procedure; continuous-time filtering in (a); and anti-alias discrete-time filtering (most often using SC techniques) in (b). Cases (c), (d) and (e) only accomplish partial channel selection because part of the filtering process is relied to digital domain.

relies on the use switched-capacitor, SC, techniques [9], [10]. Following a low-order ($n = 2$ or 3) continuous-time anti-alias filter (AAF), the switched-capacitor filter (SCF) samples the input signal and rejects the adjacent channels. Advantages of this approach are appreciable whenever the passband and stopband specifications of the AAF can be relaxed to the point that they can be met under worst-case tolerance conditions of the filter components and, therefore, tuning is not required. Most often, anti-alias filters in Fig. 14.5(b) are built using active RC techniques to preserve the high $iIP3$ achieved by the SCF. Fig. 14.6 shows three of the most usual (second order) fully-differential lowpass circuit structures used

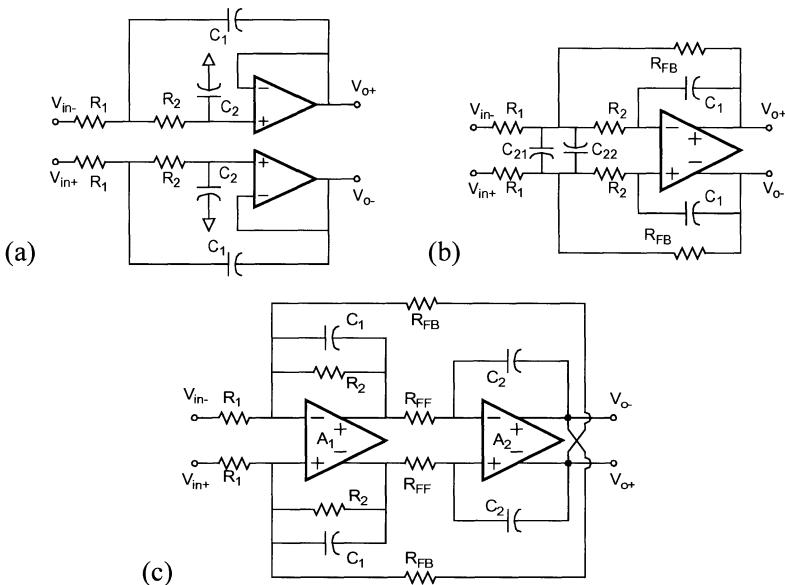


Figure 14.6: Second order active RC anti-alias filters (only a single signal path, I or Q, is shown). (a) Sallen-Key biquad; (b) Rauch biquad; (c) Tow-Thomas, or state-variable, biquad.

for anti-aliasing. In the case of complex signaling, a duplicate of the filter must be used for the quadrature path.

In order to keep the order of the anti-alias filter low, the ratio between its stopband and passband frequencies must be chosen very high, what implies that the sampling frequency, f_s , of the SCF must be much higher than the cut-off frequency, f_c , of the AAF. This can be illustrated with the help of Fig. 14.7, which shows the transfer and image characteristics for the AAF and SCF, assuming that both exhibit a lowpass response [41]. For simplicity, no sample-and-hold effect (sinc shaping) on the image response of the AAF has been considered. As shown in Fig. 14.7, if the stopband frequency of the SCF is f_b , that of the AAF can be chosen at $f_s - f_b$, because the aliased energy in the $f_b \leq f \leq f_s/2$ frequency range falls into the rejection band of the SCF.¹⁸ Thus, if the required attenuation at $f_s - f_b$ is A_{tt} , and the maximum allowed ripple in the passband is R_{pt} , the order n of a, for instance, maximally-flat, AAF is given by,

8. If no SCF were present, as in the case of Fig. 14.5(a), the stopband frequency of the continuous-time filter must be $f_s/2$ at most, to avoid the aliasing effects due to sampling at the input of the ADC.

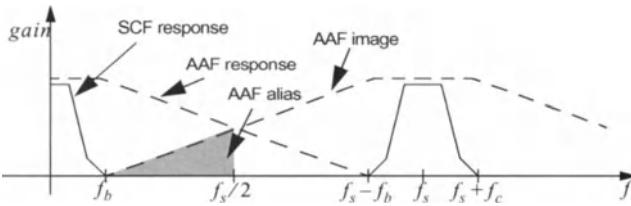


Figure 14.7: Frequency response of the anti-alias and SC filters.

$$n = \left\lceil \frac{10 \log[(A_{tt} - 1)/(R_{pl} - 1)]}{20 \log((f_s - f_b)/f_c)} \right\rceil \quad (14.4)$$

showing that the order of the AAF decreases as the ratio $(f_s - f_b)/f_c \approx f_s/f_c - 1$ is increased. However, any increment on the sampling frequency^{†9} must be traded-off, for a given cutoff frequency, with an increase on the spread of the SCF capacitances and a higher power consumption. To partially alleviate this problem, a multirate filter cascade, as shown in Fig. 14.8, is often used [10], [41]. A low-order intermediate SC filter (usually a simple biquad) is clocked at a rate N times larger than the sampling frequency, f_s , of the main (channel select) SCF. As of it, the first $N - 1$ images of the main filter lie in the stopband of the intermediate SCF, whose first image is located, in turn, well inside the stopband of the anti-alias filter. In this way, the more troublesome first image of the main SCF can be attenuated by the intermediate filter the more as the larger is f_s [10].

14.3.1.2 Mixed-signal method

In this method, which has become the preferred option in the last years, the problem of channel isolation is partitioned between the analog and digital domains. The basic principle relies on the fact that, for a given dynamic range and bandwidth, it can be found a combination of analog and digital baseband processing which is optimum in terms of power dissipation and area occupation [42]. Additional advantages are that mixed-signal partitioning features a much better

-
9. In practice, f_s can not be arbitrarily chosen but it is usually restricted by other design considerations concerning the integratability of the receiver. In a wireless scenario, for instance, f_s is usually obtained, for convenience, by dividing down the receiver's fixed LO frequency by a power of two. Another restriction is that the images of the channel-select filter lie well inside the stopband of the RF filter, whose passband is determined, in turn, by the receiver band defined by the standard (see Fig. 14.3). In a multirate cascade, to be explained below, this last restriction can be relaxed whenever any in-band image is sufficiently attenuated by the high-rate prefilter [10].

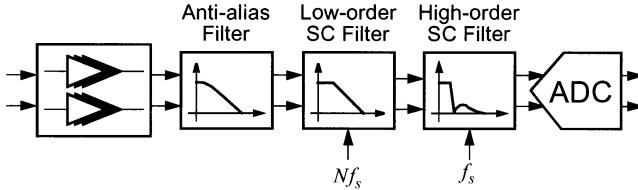


Figure 14.8: Modified SC filtering based baseband receiver section using multi-rate SCF cascade

scalability (i.e. portability for technology migration) and a great flexibility (i.e., software reconfigurable chip). Figs.14.5(c)-(e) illustrate three alternatives which mainly differ on the type of data converter used and, thereafter, on the weights by which the analog and digital filters perform the overall channel selection.

In Fig. 14.5(c), a continuous-time or sampled-data filter before a Nyquist-rate ADC performs partial channel selection in analog domain. After the ADC, a sharp transition digital FIR filter completes channel isolation. The required selectivity of the analog filter is obviously lower than in the full-analog method, however, because interferences, that may be larger than the desired signal, are not completely filtered out, the dynamic range of the ADC must be increased in consonance (8 to 10 bits are usually enough).

Another possibility for mixed-signal partitioning relies on the use of oversampled sigma-delta ($\Sigma\Delta$) modulators followed by a digital decimation filter, as shown in Fig. 14.5(d). For system level purposes, a $\Sigma\Delta$ modulator can be considered as an ADC which samples at a rate M times larger than required by aliasing considerations. For this reason, a simple low-order anti-alias filter can be used in front of the ADC. The modulator has the additional property that its quantization noise is shaped with a highpass transfer function into the frequency region above the baseband where the desired signal lies. The strong adjacent channel interferences fall into the same band as the quantization noise. This allows the same digital low-pass filter to remove both the quantization noise and adjacent channel interferences. Since this digital filter reduces the sample rate from the oversampled rate to Nyquist rate, it is referred to as a decimation filter. Because of the combined oversampling and noise shaping functions, $\Sigma\Delta$ modulators can achieve very high resolution (14 to 16 bits are not unusual) depending on the desired channel bandwidth [43]-[46].

One last mixed-signal technique for channel isolation, which is deserving considerable interest in the last years, is based on continuous-time $\Sigma\Delta$ modulation [47]-[49]. It is conceptually represented in Fig. 14.5(e). It has been shown that with this technique high resolution, high linearity ADCs can be implemented with very low power consumption. Additionally, continuous-time modulators provide

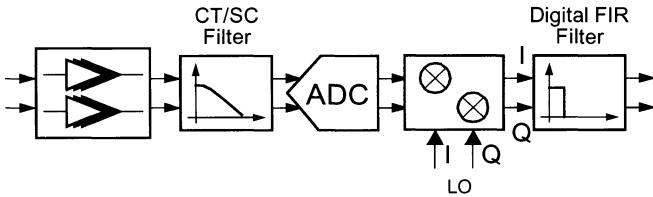


Figure 14.9: Mixed-signal baseband section for a low IF receiver –Fig. 14.3(b)– including a digital double quadrature downconverter to complete image suppression [22].

built-in anti-aliasing and less switching noise than their discrete-time counterparts. Hence, the technique represent an almost single-block solution for AFE baseband reception.

Before concluding this section, it is worth pointing out that benefits of mixed-signal partitioning is not only applicable to channel isolation, but can also extended to the problem of image rejection, as shown in Fig. 14.9 [22].

14.3.2 Signal Representation

Up to now, it has been implicitly assumed that analog baseband filtering is separately realized at both paths (I and Q) of the input signal by duplicating the same circuit topology in each branch, as shown in the left of Fig. 14.10. However, if bandpass filtering is required, a power efficient approach, which takes advantage of the quadrature relationship between the I and Q paths, can be alternatively used, the active polyphase filtering technique [22], [24].

In (bandpass) active polyphase filters, the frequency response is obtained by a linear frequency translation of a lowpass characteristic, in such a way that only positive frequency components are generated. This is in contrast with the more conventional lowpass-to-bandpass conversion based on the frequency transformation,

$$j\omega \rightarrow j\omega_c \left(\frac{\omega}{\omega_c} - \frac{\omega_c}{\omega} \right) \quad (14.5)$$

which produces passband characteristics at both $\omega = \pm\omega_c$. In active polyphase filters, however, the frequency transformation is defined as

$$j\omega \rightarrow j(\omega - \omega_c) \quad (14.6)$$

which applied, for illustration purposes, to a first order lowpass filter with transfer characteristic,

$$H_{lp}(j\omega) = \frac{1}{1 + j\omega\tau_p} \quad (14.7)$$

gives,

$$H_{bp}(j\omega) = \frac{1}{1 - j2Q + j\omega\tau_p} \quad (14.8)$$

with $Q = \omega_c\tau_p/2$ representing the quality factor of the bandpass filter, defined as the center frequency (ω_c) over the bandwidth ($2/\tau_p$). Note that if the time constant in (14.7) is defined by $\tau_p = R_p C$, as occurs in the circuit on the left of Fig. 14.11, transformation (14.6) can be simply implemented by replacing the capacitor by a composite structure with admittance $j(\omega - \omega_c)C$, as shown on top of Fig. 14.11. Given that such admittance requires quadrature signal paths, the complete bandpass filter can be directly realized as shown on the right of Fig. 14.11, where $R_c = R_p/(2Q)$.

As already mentioned, active polyphase filters not only serve the purposes of bandpass channel selection but also perform image signal rejection given that only positive frequency components are ideally allowed to pass. This property is particularly interesting on the implementation of low IF radio receiver architectures, as explained in [22], [24]. In practice, mismatches among circuit components limit the quality of the image suppression and, consequently, degrade the *BER* performance of the receiver. They are responsible of the so-called adjacent-channel leakage effect which produces a residual sideband response along with the desired signal. The power ratio of the residual output to the desired signal defines the image rejection ratio, *IRR*, which is given by [50]

$$IRR = \frac{(\Delta A/A)^2 + 2(1 + \Delta A/A)(1 - \cos\Delta\theta)}{(\Delta A/A)^2 + 2(1 + \Delta A/A)(1 + \cos\Delta\theta)} \approx \frac{(\Delta A/A)^2 + \Delta\theta^2}{4} \quad (14.9)$$

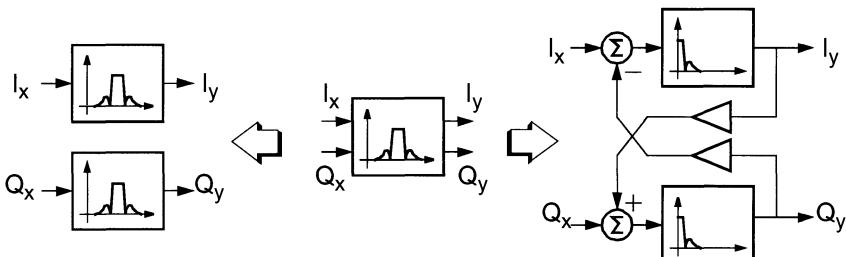


Figure 14.10: Alternative bandpass baseband filtering techniques based on real (left) and complex (right) signal processing.

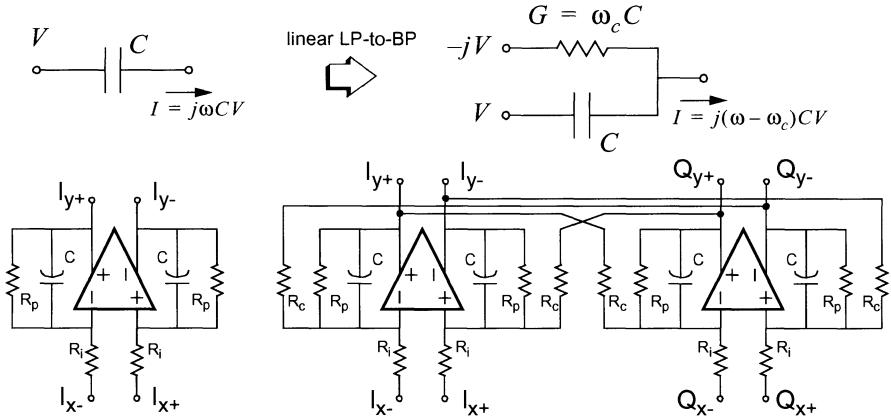


Figure 14.11: Direct synthesis of active polyphase filter from conventional low-pass prototype.

where $\Delta A/A$ and $\Delta\theta$ are, respectively, the relative gain and phase imbalances between the I and Q channels, and the approximation holds for $\Delta A \ll A$ and $\Delta\theta \ll 1\text{rad}$. Expressions for $\Delta A/A$ and $\Delta\theta$ in terms of deviations among circuit elements (for instance, resistor and capacitor mismatches in the circuits of Fig. 14.11) can be obtained through worst case analysis techniques and validated by Monte-Carlo simulations [22].

14.3.3 Filtering-Amplification Stages

In all the baseband receiver structures revised so far, the programmable gain amplifier is positioned in front of the filtering stage. This is often the situation when the filter consists of a high order structure, typically derived from a LC passive ladder prototype because of the low sensitivity to element-value variations. Such filters are, however, very noisy and the principle of amplifying before filtering is basically motivated to not substantially increase the overall noise figure of the receiver, as Friis' formula (14.1) reveals. The problem of this approach is that the linearity requirements on the filter tends to be very high because large off-channel interferers have been amplified by the same gain than the desired channel.

By interpersing amplification and filtering, the desired signal progressively grows through the baseband strip, while interferers diminish, so that at the later stages of the filter where the signal swing is largest, the interferers are substantially attenuated. The price for this improvement is that the filter must be decomposed into low-order sections what rises the sensitivity to mismatch among circuit elements. The interpersing approach is illustrated in Fig. 14.12 [51], which corresponds to a seventh order Butterworth filter used in a direct conversion receiver

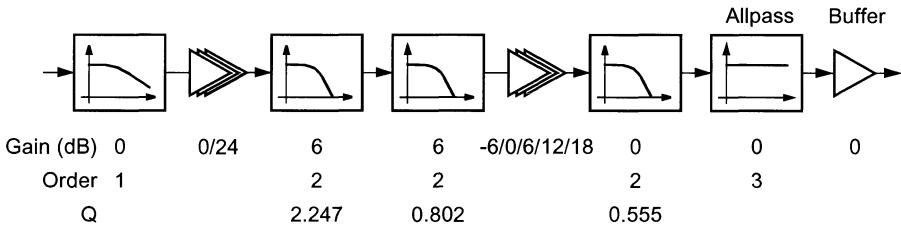


Figure 14.12: Interspersing approach for baseband amplification and filtering (only one signal path is shown).

for both DECT and DCS 1800 standards. In this example, the first stage is a passive first-order lowpass filter which provides attenuation of interferers, essentially without producing distortion. Following this single pole roll-off filter, a low input-referred noise coarse-programming amplifier is used to increase the input signal level and reduce the noise contribution of subsequent stages. Filtering operations in Fig. 14.12 are mainly concentrated in the following two biquad cells, each with an inherent gain of 6dB – they attenuate adjacent channels to a level like that of the desired signal. Then a further amplification is carried out by a second gain stage which provides five different gain levels in 6dB steps. Finally, the amplified signal is filtered again by a low-Q biquad, phase equalized by an allpass third-order filter and buffered prior to A/D conversion.

Main aspects in the interspersing approach are the gain assignment to each block in the cascade and the ordering strategy of the biquad sections. Concerning the first aspect, gain must be assigned to each block to balance signal amplification and lower noise on the one hand, and on the other to preserve good large signal handling. That is, with too high a gain the contribution of the following stages of the cascade to the overall noise figure becomes very small, but now a small input signal to the filter clips the output and compromises the overall dynamic range. Regarding the second aspect, the quality factors of the biquad sections are the main tailoring parameter to define their arrangement in the cascade. Placing high selectivity (high-Q) sections first gives more room for amplification at later stages of the cascade before output clipping occurs, what lessens their contributions to the overall noise figure. However, the noise spectrum of high-Q filter sections rises at the vicinity of the passband edges.

The interspersing approach has also interesting implications regarding power optimization. Because maximum signal levels, including the desired signal, adjacent channel blockers and out-of-band blockers, scale along the baseband strip, the slew rate requirements at the output of each block vary, what rises the possibility to optimally adjust their power consumptions [52]. The minimum slew rate

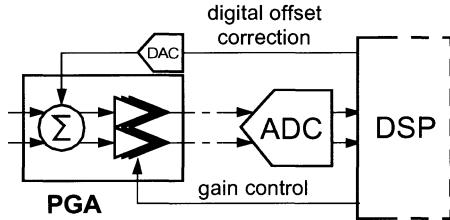


Figure 14.13: Automatic gain control (AGC) servo loop for a programmable gain amplifier.

SR_j required by the j th block of the baseband cascade to avoid signal slewing can be calculated as

$$SR_j = \max_i(2\pi f_{ij} \cdot \mathcal{V}_{ij}) \quad (14.10)$$

where \mathcal{V}_{ij} is the i th maximum signal level (desired channel or blocker) seen at the output of the j th block, and f_{ij} is the maximum frequency associated with such signal according to the blocking profile of the receiver. From (14.10), the minimum current needed in each block to slew correctly can be roughly estimated as

$$I_{SR,j} = SR_j \cdot C_j \quad (14.11)$$

where C_j is the load or compensating capacitor driven by the block.

14.3.4 Continuous/Discrete Gain Variation

Up to now, we have implicitly assumed that PGAs achieve gain variation at discrete steps (an example is shown in Fig. 14.12) by means of switch-selectable devices controlled by adjustment codes calculated in the DSP. In this case, the PGA, together with the baseband ADC and the digital circuits of the control algorithm, constitute a mixed-signal Automatic Gain Control (AGC) loop which is schematically shown in Fig. 14.13 [53]. In the most general case, the PGA also incorporates a trimming terminal to subtract any offset level induced by circuit imperfections that could be added to the signal before being amplified –this is particularly important in direct-conversion receivers where the down-converted signal has energy at dc. Offset compensation can be accomplished by means of a second servo mechanism including a DSP-controlled DAC, as depicted in Fig. 14.13, which is operated at idle times during reception. An alternative solution consists on the insertion of high-pass filters along the baseband chain [54] – the time constants of these filters must be made large enough to maintain the

received signal quality. In both cases, the objective is to avoid that the output of the PGA becomes clipped or saturates following stages of the receiver chain. Obviously, offset compensation must be efficient regardless of the adjustment code provided to the PGA and, hence, it is important that the offset level of the amplifier is not changed along with the gain steps [54].

The configuration of Fig. 14.13 offers several advantages. First, because the adjustable gain can be defined by ratios of similar devices, high accuracy amplification steps are obtained [27], [54]. Furthermore, the current consumption of the PGA can be conveniently altered at each gain setting without degrading system performance, thus leading to a net improvement on power consumption. Also, the digital circuits and baseband ADC can be shared by more than one AGC loop, in case that gain adjustment is decomposed in different stages, as previously shown.

The major drawback of PGAs arises from the discretization of the gain, particularly in modulation schemes where constellation points are very close together like 64-QAM. A sudden variation in gain can be easily mistaken for a change in signal level, and this increases the *BER* [27]. To avoid this, gain must be stepped in small increments, but this may prohibitively raise the resolution of the PGA. An efficient solution consists on partially relying amplification to a continuous variable gain amplifier (VGA) [27]. Widely linear VGAs tend to dissipate excessive power as the variation range increases, hence, a trade-off must be defined between the gain supplied by the PGAs and the VGA.

14.4 CONCLUSIONS

In this chapter, we have explored many of the design aspects which affect the implementation of active filters and PGAs in communication transceivers. Because of the widespread range of design conditions (determined by the application, transceiver architecture, and splitting of specifications among building blocks), the complexity of active filter and PGA implementations varies accordingly. In the case of filters, this goes from highly selective, low-noise implementations requiring on-chip tuning to simple untuned first order sections. The particular configuration which must be finally employed obviously depend on the fulfillment of specifications under all possible environmental conditions (desired and interfering signal levels, as well as ambient parameters such as temperature range and humidity, among others) and the reduction of area and power consumptions, particularly if the block is part of a SoC solution. Most often, circuit design is tailored from worst-case conditions (for instance, the weakest desired signal in the presence of strong interferers) which rarely occurs during reception/transmission. For this reason, most recent implementations of analog baseband sections include the ability to self-adapting under different propagation conditions in order

to provide an average saving of power consumption [27].

APPENDIX 14.I NOISE FIGURE [19], [20], [50], [63]

The noise figure NF is a quantitative measure of how much the signal-to-noise ratio SNR degrades as the signal passes through a circuit block or a system. It depends on several factors such as losses in the circuit, the solid-state devices, bias applied, and amplification. The noise factor of a two-port network is defined as^{†10}

$$F = \frac{p_s/n_s}{p_o/n_o} = \frac{snr_s}{snr_o} \quad (14.I.1)$$

where p and n represent the available signal and noise powers^{†11}, and the subscripts s and o represent the driving source and output of the network, respectively. The noise figure NF is simply the noise factor converted in dB notation. If the available power gain of the block is g_a (defined in (14.I.3)), the noise factor can be also expressed as

$$F = \frac{p_s(n_s g_a + n_o)}{n_s p_s g_a} = 1 + \frac{n_o}{n_s g_a} = 1 + \frac{n_i}{n_s} \quad (14.I.2)$$

where n_o and n_i are, respectively, the output- and input-referred available noise powers intrinsically generated by the block, and n_s is the available source noise power. Hence, the noise factor can be alternatively defined as the ratio of the total

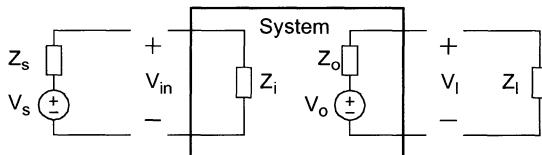


Figure 14.14: Generic time-invariant system and associated impedances in a driven ad loaded configuration.

-
- 10. As a general notation rule, in the case of power quantities (either referring to power levels, gains, or ratios) we will use lower case letters to denote linear values and upper case letters to denote decibel values. Thus, for instance, p_s watts corresponds to $P_s = 10\log(p_s)$ [dB]. The only exception to this rule is made with respect to the noise figure measure, following most extended nomenclature.
 - 11. The available power p_x from a voltage source V_x with internal impedance Z_x , is the power that the source would deliver to a conjugated matched circuit, Z_x^* (maximum power transfer condition). If R_x is the real part of Z_x , the available power is then found to be $p_x = V_x^2 / (4R_x)$.

available input-referred noise power ($n_s + n_i$) to that of the source alone. On the other hand, if the network, assumed time-invariant with load impedance Z_l , has an input impedance Z_i and an output impedance Z_o , the available power gain is defined as

$$g_a = \left| \frac{Z_i}{Z_i + Z_s} \right|^2 |a_v|^2 \frac{R_s}{R_o} \equiv \alpha^2 |a_v|^2 \frac{R_s}{R_o} \quad (14.I.3)$$

where $a_v = V_o / V_{in}$ is called the (unloaded) voltage gain and Z_s is the internal impedance of the driving voltage source V_s , as depicted in Fig. 14.14. R_s and R_o are resistances defined by the real part of impedances Z_s and Z_o , respectively. Interestingly, in the case of passive networks, for which the total output-referred available noise power ($n_s g_a + n_o$) coincides with available noise power at the source (n_s), the noise factor simplifies to $F = 1/g_a \equiv l_a$, where l_a is referred to as the insertion loss of the network.

Assuming that the observation bandwidth B_s is low enough such that resistances do not vary substantially with frequency¹², the mean-square thermal noise voltage generated by Z_s at a given frequency can be found to be $v_{ns}^2 = 4kT R_s B_s$, where k is the Boltzmann's constant, T is the absolute temperature. Hence, $n_s = v_{ns}^2 / (4R_s) = kTB_s$ and replacing this in the last equality of (14.I.2) obtains,

$$F = 1 + \frac{n_i}{kTB_s} \quad (14.I.4)$$

On the other hand, n_i can be expressed in terms of the input-referred noise sources v_{ni}^2 and i_{ni}^2 as $n_i = (v_{ni} + Z_s i_{ni})^2 / (4R_s)$ (see Fig. 14.15) and therefore

$$F = 1 + \frac{(v_{ni} + Z_s i_{ni})^2}{4kT R_s B_s} \quad (14.I.5)$$

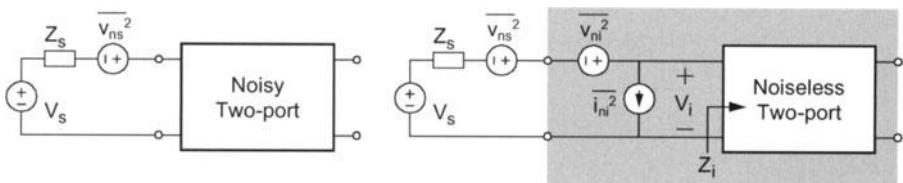
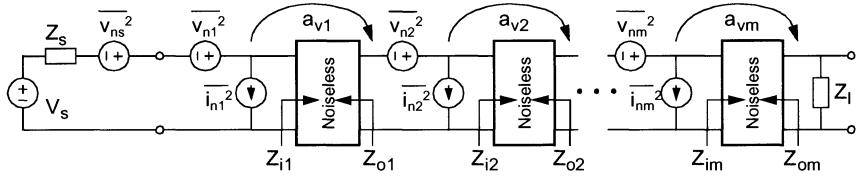


Figure 14.15: Representation of noise in a two-port network by equivalent input noise generators.

12. Conventionally, B_s is chosen as 1Hz, and NF is then called spot noise figure.

Figure 14.16: Cascade of m noisy stages.

where it should be noted that v_{ni} and $Z_s i_{ni}$ (measured in the same observation bandwidth B_s and at the same frequency) have been added before squaring to account for their correlation. An identical result could be obtained from the previous equality of (14.I.2), by expressing n_o in terms of the intrinsic output referred thermal noise power generated by the network, v_{no}^2 , as $n_o = v_{no}^2 / (4R_o)$, and noting that $v_{no}^2 = \alpha^2 |a_v|^2 (v_{ni} + Z_s i_{ni})^2$.

For a cascade of two-port networks, the overall noise factor can be obtained in terms of the F_j and available power gain g_{aj} of each stage. They are related by the general expression (assuming m stages),

$$F_{tot} = 1 + (F_1 - 1) + \frac{(F_2 - 1)}{g_{a1}} + \dots + \frac{(F_m - 1)}{g_{a1} \dots g_{a, m-1}} \quad (14.I.6)$$

which is known as the Friis equation. It must be emphasized that this equation is only correct when available power gains are used. Interestingly, if the P first elements of the cascade are passive, as occurs in wireless receivers where the front-end of the RX medium interface contains at least an antenna filter, (14.I.6) can be written as,

$$F_{tot} = l_{a1} \dots l_{aP} \left[1 + (F_{P+1} - 1) + \dots + \frac{(F_m - 1)}{g_{a, P+1} \dots g_{a, m-1}} \right] \quad (14.I.7)$$

where l_{a1}, \dots, l_{aP} are the insertion losses of the first P (passive) blocks of the cascade.

Taking into account (14.I.3) and (14.I.5), and using the notation of Fig. 14.16, we can write (14.I.6) as

$$F_{tot} = 1 + \left(\frac{1}{4KTR_s B_s} \right) \sum_{j=1, m} \frac{(v_{nj} + Z_{o, j-1} i_{nj})^2}{\prod_{k=1, j} \alpha_{k-1}^2 a_{v, k-1}^2} \quad (14.I.8)$$

where $\alpha_k = |Z_{ik}/(Z_{ik} + Z_{o, k-1})|$, with $Z_{o, 0} \equiv Z_s$ and $\alpha_0 = a_{v0} = 1$. Interestingly, if the input impedance of a block is matched to the output impedance of the

previous one ($Z_{ik} = Z_{o,k-1}$) we have $\alpha_k = 1/2$, whereas if the input impedance of a block is much higher than its source impedance, $\alpha_k = 1$.

Often equation (14.I.8) is simplified by defining equivalent noise resistances R_{ej} for each of the stages in the cascade, whose thermal noise power in bandwidth B_s , $v_{ej}^2 = 4KTR_{ej}B_s$, coincides with $(v_{nj} + Z_{o,j-1}i_{nj})^2$. In this case, [55]

$$F_{tot} = 1 + \sum_{j=1,m} \frac{R_{ej}/R_s}{\prod_{k=1,j} \alpha_{k-1}^2 a_{v,k-1}^2} \quad (14.I.9)$$

If we assume that all the blocks in the cascade are built on-chip and verify that $Z_{ik} \gg Z_{o,k-1}$ with the exception of the first block, whose input impedance is matched to the off-chip source resistance R_s (conventionally, 50Ω), then (14.I.9) can be expressed as,

$$F_{tot, integrated} = \frac{R_{board} + R_{e1}/4 + \sum_{j=2,m} R_{ej}/\prod_{k=1,j} a_{v,k-1}^2}{R_{board}} \quad (14.I.10)$$

$$\equiv \frac{R_{board} + R_{integrated}}{R_{board}}$$

where $R_{board} = R_s/4$. If the blocks $P+1$ to m in (14.I.7) are assumed to be fully integrated (under conditions above) so that they can be described by an equivalent noise resistance, $R_{integrated}$, in the terms defined in (14.I.10), then the overall noise factor of the cascaded network, including the first P passive elements is given by,

$$F_{tot} = l_{a1} \cdots l_{aP} \left(1 + \frac{R_{integrated}}{R_{board}} \right) \quad (14.I.11)$$

14.I.1 Sensitivity

The lowest signal level that a system is able to detect with an acceptable carrier-to-noise-and-interference ratio, $C/(I+N)_{min}$, at its output, is called the minimum detectable signal (*MDS*) level, and it is often used as a reference for the sensitivity of a system. The *MDS* is calculated by:

$$MDS = N_r + NF_{tot} + C/(I+N)_{min} \quad (14.III.1)$$

where N_r is the available thermal noise power at the system input^{†13}, NF_{tot} is the total noise figure of the system (discussed in Appendix 14.I), and $C/(I+N)_{min}$ is

the minimum signal-to-noise ratio required to maintain a certain level of signal processing quality, for instance, a prescribed bit error rate (*BER*). For GSM, the reference sensitivity level is required to be -102dBm [14], and taking into account that a minimum $C/(I+N)$ of 9dB must be guaranteed, it can be found from (14.III.1) that the maximum cumulated noise figure of the receiver measured at the antenna input is 9.85dB .

Another important measure is the noise floor defined as the total integrated noise available at the input of the system and given by

$$N_{\text{floor}} = N_r + NF_{\text{tot}} \quad (14.\text{III}.2)$$

which is -111dBm for the GSM standard.

APPENDIX 14.II LINEARITY MEASURES [19], [50]

The linearity of a receiver determines the maximum allowable input signal level, before distortion generated by circuit components significantly degrades communication performance. To analyze and quantify distortion effects, let's consider a memoryless, time-invariant network whose input-output characteristic can be modeled, in general, by the polynomial expansion,

$$V_o = a_0 + a_1 V_{\text{in}} + a_2 V_{\text{in}}^2 + a_3 V_{\text{in}}^3 + \dots \quad (14.\text{II}.1)$$

and suppose that a two-tone input signal,

$$V_{\text{in}} = V_1 \cos(\omega_1 t) + V_2 \cos(\omega_2 t) \quad (14.\text{II}.2)$$

is applied to the system. Further, assume that ω_1 and ω_2 are close enough so that coefficients a_k can be considered unchanged at both tones. If the network is weakly nonlinear in the sense that coefficients higher than cubic in (14.II.1) can be considered negligible for most practical purposes, the response will contain spectral components at the frequencies listed in Table 14.II.1 and represented in Fig. 14.17.

Note that the nonlinear (distortion) terms, a_2 and a_3 , in the input-output transfer characteristic of the network are responsible for the generation of new signals at frequencies other than the fundamental tones, and make that the theoretically constant circuit gain, a_1 , becomes dependent on the applied input level. These deviations from linearity manifest in a number of effects such as gain compres-

13. Assuming thermal equilibrium, the available thermal noise power is given by $n_r = k \cdot T \cdot B_w$, where B_w is the bandwidth of interest. In the GSM standard, channels spacing is 200kHz wide, thus, $N_r = -173.86\text{dBm/Hz} + 10\log(200\text{kHz}) = -120.85\text{dBm}$, assuming $T = 298\text{K}$.

Table 14.II.1: Frequency components generated by a two-tone input.

	Frequency	Amplitude, V_o
DC	0	$a_0 + \frac{1}{2}a_2(V_1^2 + V_2^2)$
Fundamental	ω_1	$[a_1 + a_3\left(\frac{3}{4}V_1^2 + \frac{3}{2}V_2^2\right)]V_1$
	ω_2	$[a_1 + a_3\left(\frac{3}{2}V_1^2 + \frac{3}{4}V_2^2\right)]V_2$
Second Order Terms	$2\omega_1$	$\frac{1}{2}a_2V_1^2$
	$2\omega_2$	$\frac{1}{2}a_2V_2^2$
	$\omega_1 \pm \omega_2$	$a_2V_1V_2$
Third Order Terms	$3\omega_1$	$\frac{1}{4}a_3V_1^3$
	$3\omega_2$	$\frac{1}{4}a_3V_2^3$
	$2\omega_1 \pm \omega_2$	$\frac{3}{4}a_3V_1^2V_2$
	$\omega_1 \pm 2\omega_2$	$\frac{3}{4}a_3V_1V_2^2$

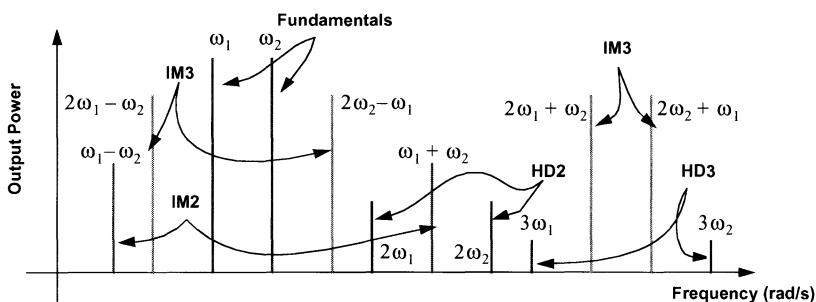


Figure 14.17: Effects of distortion on the output spectrum.

sion, blocking and intermodulation which are briefly discussed below.

14.II.1 Gain Compression

Gain compression^{†14} occurs when a network cannot increase its output amplitude in linear proportion to an increase in the input amplitude. This effect can be explained by observing the output amplitudes of the fundamental terms in Table 14.II.1. If the sign of coefficient a_1 in (14.II.1) is opposite to that of coefficient a_3 , as occurs in most practical situations, the ideal gain of the network, a_1 , is decremented by an amount proportional to a weighted sum of the power of the two input tones.

Gain compression is characterized by the so-called 1-dB (input-referred) compression point, iCP ^{†15}, which is defined as the input power at which the fundamental tone at the output, P_o , drops by 1 dB with respect to the ideal linear response. It, thus, can be calculated from the expression,

$$[P_o - (P_{in} + G)] \Big|_{P_{in} = iCP} = -1 \text{ dB} \quad (14.II.3)$$

where G is the (small-signal) power gain of the network^{†16}. Taking into account that $p_{in} = 0.5V_I^2R_i$ and $p_o = 0.5V_O^2R_l/(R_l + R_o)^2$, where V_I and V_O are, respectively, the amplitude of the fundamental tones at the input and output of the system, the above equation can be expressed in linear notation as

$$(V_O'/V_I) \Big|_{V_I = V_{ICP}} = |a_1|(10^{-1/20}) \quad (14.II.4)$$

where $V_{ICP} = \sqrt{2R_i icp}$ is the input level of the 1-dB compression point. Equation (14.II.4) can be further evaluated taking into account, the (distorted) output amplitude at the fundamental frequency of Table 14.II.1. If a single input tone is present at the input of the network ($V_1 = V_I$ and $V_2 = 0$), (14.II.4) gives,

$$icp = \frac{2}{3R_i} \left| \frac{a_1}{a_3} \right| (1 - 10^{-1/20}) \quad (14.II.5)$$

which reveals that a smaller $|a_3|$ results in lower distortion and, therefore, in a higher compression point. In case that two equal tones are presented to the net-

14. This must be distinguished from the gain saturation effect which occurs when the output amplitude stops increasing (or even decrease) for increasing input amplitude.
15. In general, power levels referred to the input (output) of a system will be denoted with a lowercase 'i' ('o') preceding the name of the quantity.
16. Referring back to Fig. 14.14 and assuming that impedances can be treated as resistors, the power gain of a network is defined as ratio between the powers delivered at the input $P_{i,del} = V_I^2/R_i$ and the load $P_{l,del} = V_o^2R_l/(R_l + R_o)^2$. Hence, $g = |a_v|^2R_lR_i/(R_l + R_o)^2$ where a_v is the small-signal (unloaded) voltage gain of the system [63] (for the system described by (14.II.1) we have $a_v = a_1$).

work, it can be shown that icp is three times lower than the value in (14.II.5).

Gain saturation can be also analyzed as an average gain compression effect, leading to an alternative definition of the 1-dB compression point. It can be shown that the rms value of a sinewave falls by 1 dB when it is symmetrically clipped to about 79% of its undistorted amplitude [19]. Hence, if we denote as $V_{o,sat}$ the output saturation level of the circuit, the 1-dB compression point due to clipping is given by,

$$icp = \frac{1}{2R_i} \left(\frac{V_{o,sat}}{0.79a_1} \right)^2 \quad (14.II.6)$$

where it is assumed that no third order effect is present. Equation (14.II.6) can be expressed in decibel notation as $iCP \approx iP_{sat} + 2\text{dB}$, where iP_{sat} is the maximum output power capability of the network referred to its input – usually an additional error budget is included in the above expression to account with the neglected third order effects. In practice, the lowest of the two values in (14.II.5) and (14.II.6) is taken as a measure of the maximum input range of the network and, hence, a reference to define the upper bound of the dynamic range, as will be seen in Appendix 14.III.

If the network is composed by the cascade of several two-ports, it can be shown that the overall compression point can be obtained from the compression points of its building blocks, through the equation [50]

$$\frac{1}{icp_{tot}} \approx \frac{1}{icp_1} + \frac{g_1}{icp_2} + \frac{g_1g_2}{icp_3} + \dots \quad (14.II.7)$$

where g_j is the (uncompressed) power gain of the $j - th$ element in the chain.

14.II.2 Blocking and Desensitization

These effects are also related to the compressive input-output characteristics exhibited by most circuits of interest, and occurs when a weak desired signal is processed along with a strong interferer. The desensitization effect consists on the reduction of the small-signal gain experienced by the weak signal as a result of interference. If interference is strong enough, gain may even drops to zero and then, we say that the desired weak signal is blocked by the interferer. To see this effect, let's assume $V_1 \ll V_2$ and consider the fundamental term at ω_1 . The ratio of the distorted to the ideal (unloaded) voltage gain of the network can be approximated as

$$\rho \approx 1 + \frac{3}{2} \left(\frac{a_3}{a_1} \right) V_2^2 \quad (14.II.8)$$

and, bearing in mind that a_1 and a_3 have opposite signs, blocking is observed when V_2 is large enough such that $\rho_2 = 0$. A common way to quantify desensitization is by finding the interferer input power required to reduce the voltage gain at the fundamental tone by 3-dB ($20\log(\rho) = -3\text{dB}$). It can be calculated as

$$idp = \frac{1}{3R_i} \left| \frac{a_1}{a_3} \right| (1 - 10^{-3/20}) \quad (14.II.9)$$

which, compared with (14.II.5), shows that the 3-dB desensitization point, iDP , is about 1.2dB above the 1-dB compression point, iCP .

14.II.3 Harmonics Generation

Among the spurious responses generated by the system of (14.II.1), an important set is formed by those spectral components which appear at integer multiples of the fundamental frequencies. They are called harmonic responses and arise even if a single-tone is applied to the network. In this case and according to Table 14.II.1, the output level of the $q - th$ harmonic response is given by

$$V_{HD, q} = \frac{1}{2^{q-1}} a_q V_I^q \quad (14.II.10)$$

where V_I is the signal level of the applied tone^{†17}. The $q - th$ harmonic response is typically characterized by the $q - th$ harmonic distortion factor, HD_q , which is defined as the difference between the powers, in dB notation, of the (undistorted) fundamental and the $q - th$ harmonic response, H_q , both measured at the output of the system, i.e., $HD_q = P_{o, ideal} - H_q$ where $P_{o, ideal} = P_{in} + G$ and G is the (small-signal) power gain of the network. Using (14.II.10), the $q - th$ harmonic distortion factor can, then, be expressed in linear notation as

$$hd_q = \left[\frac{a_1}{a_q} \left(\frac{2}{V_I} \right)^{q-1} \right]^2 = \left[\frac{a_1}{a_q} \right]^2 \left(\frac{2}{R_i p_{in}} \right)^{q-1} \quad (14.II.11)$$

The total harmonic distortion is defined as the sum of all the harmonic distortion factors $thd = hd_2 + hd_3 + \dots$ (in practice, this is restricted to the most significant factors, usually from 2nd to 5th) or, in decibel notation,

17. This expression is only an approximation if terms with order higher than cubic are considered in (14.II.1).

$$THD = 10\log(thd).$$

14.II.4 Intermodulation

When a two-tone input signal like (14.II.2) is fed into a nonlinear network, its output will contain spurious responses at frequencies $j\omega_1 \pm k\omega_2$, $j, k = 1, 2, \dots$. Such responses are referred to as intermodulation (IM) products of order $q = j + k$ and can profoundly affect the performance even of systems operated far below gain compression. For the case described in Table 14.II.1, the second-order IMD products appear at $\omega_1 \pm \omega_2$, and those of third-order are the signals at $2\omega_1 \pm \omega_2$ and $\omega_1 \pm 2\omega_2$.

Generation of IM products are conventionally quantified by the so-called intermodulation intercept points which can be derived from a two-tone test. This test consists on applying two tones with the same power p_{in} to the network, and evaluate the increments of the IM products as p_{in} increases. Under small-signal conditions (at levels well below compression), the power of the $q - th$ order IM product is proportional to the test input power rised to q – this can be understood from Table 14.II.1 noting that the magnitude of the $q - th$ IM product grows in proportion to V_I^q , where $V_I = V_1 = V_2$ is the signal level of the applied tones. On a log-log representation of the (power) transfer characteristics of the system, this implies that the $q - th$ order IM product grows at q times the rate at which the fundamental component increases. The (input-referred) $q - th$ order intercept point, iIP_q , is then defined to be the intersection of the two lines, i.e., the input power (of one tone) at which the power of the $q - th$ order IM product would be equal to that of the fundamental (if compression did not occur). In practice, because gain compression effects become noticeable before intercept points are reached, they must be determined by extrapolation of both the fundamental and the IM product – this is illustrated in Fig. 14.18 (powers are expressed in dBms) for the second and third intercept points. Taking into account the above definition, the line which describes the variation of the (output-referred) power, oIM_q , (or simply IM_q) of the $q - th$ order IM product with the input power tone is given by,

$$IM_q - oIP_q = q(P_{in} - iIP_q) \quad (14.II.12)$$

where oIP_q is the (output-referred) $q - th$ intercept point. Since intercept points also belong to the (extrapolated) line of the fundamental output tone, we have

$$oIP_q = iIP_q + G \quad (14.II.13)$$

where G is the (small-signal) power gain of the network. Combining (14.II.12) and (14.II.13), we obtain

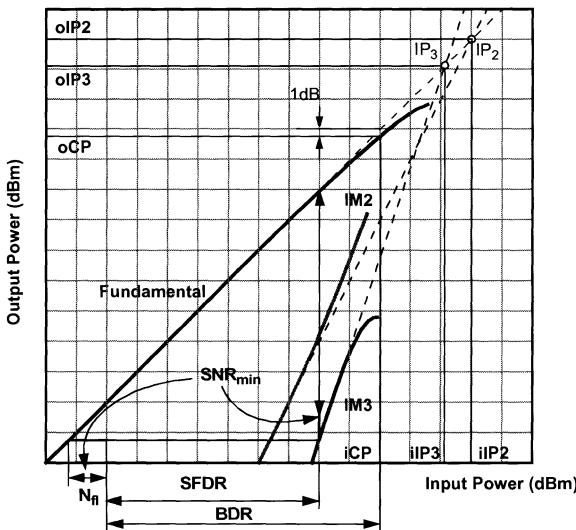


Figure 14.18: Graphical interpretation of the second- and third-order intercept points and the 1-dB compression point. Each curve depicts the power of the spectral component evaluated. All quantities (iCP , iIP_2 , iIP_3 and corresponding output-referred values oCP , oIP_2 , oIP_3) are expressed in dBm units.

$$iIP_q = \frac{1}{(q-1)}[qP_{in} - IM_q + G] \equiv P_{in} + \frac{\Delta P_q}{(q-1)} \quad (14.II.14)$$

where ΔP_q is the difference between the powers, in decibels, of the (undistorted) fundamental and the $q-th$ IM product at the output, $\Delta P_q = P_{o,ideal} - IM_q$ where $P_{o,ideal} = P_{in} + G$. Note that ΔP_q can be interpreted as a signal-to-distortion ratio which can be also evaluated at the input of the system as $\Delta P_q = P_{in} - iIM_q$, where iIM_q is the input-referred $q-th$ IM product. Taking into account that $p_{in} = 0.5 V_I^2 R_i$, the above equation can be expressed in linear notation as

$$iip_q = \frac{1}{2R_i} \left[\frac{a_1 V_I^q}{V_{IM,q}} \right]^{\frac{2}{q-1}} \quad (14.II.15)$$

where $V_{IM,q}$ is the signal level associated to the $q-th$ IM product. Assuming equal amplitude for the input tones, we have from Table 14.II.1 that $V_{IM,q} = C_{IM,q} a_q V_I^q$, where $C_{IM,q}$ is a real coefficient (see footnote 17) and, hence, iip_q can be expressed in terms of the coefficients in the polynomial expan-

sion (14.II.1) as,

$$iip_q = \frac{1}{2R_i} \left[\frac{1}{C_{IM,q}^2} \left(\frac{a_1}{a_q} \right)^2 \right]^{\frac{1}{q-1}} \quad (14.II.16)$$

Thus, for instance, in case of the 2nd order intercept point, we have $C_{IM,2} = 1$ and, therefore,

$$iip_2 = \frac{1}{2R_i} \left(\frac{a_1}{a_2} \right)^2 \quad (14.II.17)$$

and, for the 3rd order intercept point, we have $C_{IM,3} = 3/4$ and, hence,

$$iip_3 = \frac{2}{3R_i} \left| \frac{a_1}{a_3} \right| \quad (14.II.18)$$

which, compared with (14.II.5), it can be seen that the difference between the 1-dB compression point and the input-referred 3rd order intercept point is at least -9.6dB .

Interestingly, intercept points are related to the harmonic distortion factors of the same order through the expression (see (14.II.11)),

$$iip_q = \frac{p_{in} h d_q^{1/(q-1)}}{(2C_{IM,q})^2} \quad (14.II.19)$$

or in decibel notation,

$$iIP_q = P_{in} + \frac{HD_q}{q-1} - 20\log(2C_{IM,q}) \quad (14.II.20)$$

as can be deduced from (14.II.11) and (14.II.15).

The overall intercept points of a system formed by the cascade of several two-ports can be also expressed in terms of corresponding figures of its building blocks. Focusing on the most relevant in practice, second- and third-order intercept points, we have, respectively [50]

$$\frac{1}{\sqrt{iip_{2,tot}}} \approx \frac{1}{\sqrt{iip_{2,1}}} + \frac{\sqrt{g_1}}{\sqrt{iip_{2,2}}} + \frac{\sqrt{g_1 g_2}}{\sqrt{iip_{2,3}}} + \dots \quad (14.II.21)$$

and

$$\frac{1}{iip_{3,tot}} \approx \frac{1}{iip_{3,1}} + \frac{g_1}{iip_{3,2}} + \frac{g_1 g_2}{iip_{3,3}} + \dots \quad (14.II.22)$$

where $iip_{2,j}$ and $iip_{3,j}$ are, respectively, the input-referred second- and third-order intercept points of the $j - th$ stage, and g_j is the (uncompressed) power gain of the $j - th$ stage. It must be emphasized that the above expressions are approximations which assume the worst-case condition in that all distortion products add in-phase[20].

14.II.5 Distortion Induced Offset

According to Table 14.II.1, the second-order distortion term, a_2 , in (14.II.1) is responsible of introducing an additional DC component to the network's response. Assuming a single-tone input signal, the input-referred power of this extra DC contribution is given by,

$$ip_{DC} = \frac{1}{R_i} \left(\frac{1}{2} \frac{a_2}{a_1} V_I^2 \right)^2 = \frac{1}{2} \frac{P_{in}^2}{iip_2} \quad (14.II.23)$$

or

$$iP_{DC} \approx 2P_{in} - iP_2 - 3\text{dB} \quad (14.II.24)$$

This distortion induced offset is particularly troublesome in systems which translate modulated information to DC, as occurs in zero-IF receivers, putting severe demands on the required iP_2 .

APPENDIX 14.III DYNAMIC RANGE

The dynamic range of a system is roughly defined as the signal range where the output power is almost linearly proportional to its input power. The higher the dynamic range is, the more tolerant the receiver is towards signals that vary greatly in magnitude. Particular definitions of dynamic range depends on how the lower and upper bounds are assigned.

For the blocking dynamic range (BDR), the lower bound is defined by the minimum detectable signal and the upper bound by the 1-dB compression point (see Appendix 14.II). Hence,

$$BDR = iCP - MDS \quad (14.III.3)$$

where iCP must be higher than the maximum expected input signal to the system. For instance, in GSM the maximum specified out-of-band blocker has a power of

+0dBm, therefore, defining $iCP = +3\text{dBm}$ to avoid the saturation onset (see Appendix 14.II), the minimum BDR required at the antenna input is 102dB, taking into account the MDS value imposed by the standard (see Appendix 14.I).

In other cases, the upper bound of the dynamic range is defined by the input power at which the total harmonic distortion reaches a limit value of, say, 60dB. This definition is the most familiar among analog integrated circuit designers.

Another definition of dynamic range is the spurious-free dynamic range ($SFDR$) which also differs from BDR on the upper bound set. Such bound is defined as the maximum input level in a two-tone test for which the IMD products referred to the input of the system do not exceed the noise floor, i.e., $iIM_q - G < N_{floor}$ for $q = 2, 3, \dots$ (see Appendix 14.II). From (14.II.14), the maximum input power is, then, found to be

$$P_{in, max} = \min_q \left[\frac{(q-1)iIP_q + N_{floor}}{q} \right] \quad (14.III.4)$$

where iIP_q is the input-referred $q-th$ intercept point. The $SFDR$ is the difference between $P_{in, max}$ and the minimum detectable signal, and hence,

$$\begin{aligned} SFDR &= P_{in, max} - MDS \\ &= \min_q \left[\frac{(q-1)}{q} (iIP_q - N_{floor}) \right] - C/(I+N)_{min} \end{aligned} \quad (14.III.5)$$

In the case of wireless receivers, the minimum required spurious-free dynamic range must be deduced from specifications and tests defined in the physical layer recommendations of the targeted standard. In the particular case of GSM, such minimum requirement is found through the intermodulation test and the maximum input power level specification [64].

The intermodulation test specifies that a GSM signal 3dB above the reference sensitivity level ($P_{min} = MDS + 3\text{dB} = -99\text{dBm}$, see Appendix 14.I) must be detectable ($C/(I+N) \geq 9\text{dB}$) in the presence of two interferers with power $P_{test} = -49\text{dBm}$: a static sinewave and a GMSK modulated signal, located at 800kHz and 1.6MHz away from the desired channel, respectively [14]. Assuming that the interference and noise components contribute equally to the ratio $C/(I+N)$, we have $C/N = C/I \geq 12\text{dB}$ and, therefore, the input-referred noise floor in this test set-up must satisfy $N_{floor} < P_{min} - C/N = -111\text{dBm}$ [55]. Similarly, the input-referred $q-th$ IM product must satisfy the condition $iIM_q < P_{min} - C/I = -111\text{dBm}$. Hence, from (14.II.14) we have,

$$iIP_q = \frac{qP_{test} - iIM_q}{q-1} > \frac{qP_{test} - (P_{min} - C/I)}{q-1} \quad (14.III.6)$$

which, for the 2nd and 3rd order intercept points gives, $iIP_2 > 13\text{dBm}$ and $iIP_3 > -18\text{dBm}$, respectively.

On the other hand, harmonic distortion components must be always remain below the wanted signal level by a minimum C/I amount to guarantee a correct reception. The worst-case situation occurs when the input signal reaches its maximum power level, P_{max} , which for GSM amounts -15dBm . Hence, condition

$$HD_q|_{P_{in} = P_{max}} > C/I \quad (14.\text{III}.7)$$

must be met. Given that harmonic distortion factors and intermodulation products are related between themselves, as demonstrated in Appendix 14.II, the above condition can be used to derive corresponding expressions for the intercept points. Namely, for the 2nd intercept point, we have from (14.II.20)

$$iIP_2 > P_{max} + C/I - 6\text{dB} \quad (14.\text{III}.8)$$

which gives $iIP_2 > -9\text{dBm}$ for $C/I \geq 12\text{dB}$, a weaker condition than that obtained by the intermodulation test. Similarly, for the 3rd intercept point,

$$iIP_3 > P_{max} + (1/2)C/I - 3.5\text{dB} \quad (14.\text{III}.9)$$

and, therefore, $iIP_3 > -12.5\text{dBm}$ which is larger than that derived from (14.III.6). Taking this last value into (14.III.5), we finally obtain that $SFDR$ must be larger than 57dB .

14.5 REFERENCES

- [1] C. Chien, *Digital Radio Systems on a Chip. A Systems Approach*. Kluwer Academic Publishers, Norwell (MA), 2001.
- [2] A. A. Abidi, P. R. Gray and R. G. Meyer (Eds.), *Integrated Circuits for Wireless Communications*. IEEE Press, New York, 1999.
- [3] Y. P. Tsividis, and J. Voorman, *Integrated Continuous-Time Filters: Principles, Design and Applications*. IEEE Press, New York, 1992.
- [4] B. Nauta, *Analog CMOS Filters for Very High Frequencies*. Kluwer Academic Publishers, Boston (MA), 1992.
- [5] S. D. Willingham, and K. W. Martin, *Integrated Video-Frequency Continuous-Time Filters. High-Performance Realizations in BiCMOS*. Kluwer Academic Publishers, Boston (MA), 1995.
- [6] S. Pavan, and Y. P. Tsividis, *High Frequency Continuous Time Filters in Digital CMOS Processes*. Kluwer Academic Publishers, Boston (MA), 2000.
- [7] A. Sedra, and P. Brackett, *Filter Theory and Design: Active and Passive*. Matrix Publishers, Beaverton (OR), 1968.

- [8] R. Schaumann, K. Laker, and M. Ghausi, *Analog Filter Design – Passive, Active RC and Switched-Capacitor*. Prentice Hall Publishers, Englewood-Cliffs (NJ), 1992.
- [9] J. C. Rudell, J.-J. Ou, T. B. Cho, G. Chien, F. Brianti, J. A. Weldon, and P. R. Gray, “A 1.9-GHz Wide-Band IF Double Conversion CMOS Receiver for Cordless Telephone Applications,” *IEEE J. of Solid-State Circuits*, vol. 32, pp. 2071–2088, December 1997.
- [10] P. J. Chang, A. Rofougaran, and A. A. Abidi, “A CMOS Channel-Select Filter for a Direct-Conversion Wireless Receiver,” *IEEE J. Solid-State Circuits*, vol. 32, pp. 722–729, May 1997.
- [11] M. Gustavsson, J. J. Wikner and N. N. Tan, *CMOS Data Converters for Communications*. Kluwer Academic Publishers, Boston (MA), 2000.
- [12] P. P. Siniscalchi, J. K. Pitz, R. K. Hester, S. M. DeSoto, M. Wang, S. Sridharan, R. L. Halbach, D. Richardson, W. Bright, M. M. Sarraj, J. R. Hellums, C. L. Betty, and G. Westphal, “A CMOS ADSL Codec for Central Office Applications,” *IEEE J. Solid-State Circuits*, vol. 36, pp. 356–365, March 2001.
- [13] D. S. Langford, B. J. Tesch, B. E. Williams, and G. R. Nelson, “A BiCMOS Analog Front-End Circuit for an FDM-Based ADSL System,” *IEEE J. Solid-State Circuits*, vol. 33, pp. 1383–1393, September 1998.
- [14] “Digital cellular telecommunication system: Radio transmission and reception (GSM 5.05),” European Telecommunication Standards Inst. (ETSI), Sophia Antipolis, France.
- [15] T. Stetzler, I. Post, J. Havens, and M. Koyama, “A 2.7–4.5 V single chip GSM transceiver RF integrated circuit,” *IEEE J. Solid-State Circuits*, vol. 30, pp. 1421–1429, December 1995.
- [16] T. Yamawaki, M. Kokubo, K. Irie, H. Matsui, T. Endou, H. Hagiwara, T. Furuya, Y. Shimizu, M. Katagishi, and J. Hildersley, “A 2.7-V GSM RF transceiver IC,” *IEEE J. Solid-State Circuits*, vol. 32, pp. 2089–2096, December 1997.
- [17] P. Orsatti, F. Piazza, and Q. Huang, “A 20-mA-receive, 55-m A-transmit, single-chip GSM transceiver in 0.25- μ m CMOS,” *IEEE J. Solid-State Circuits*, vol. 34, pp. 232–233, December 1999.
- [18] A. A. Abidi, “Direct conversion radio transceivers for digital communications,” *IEEE J. Solid-State Circuits*, vol. 30, pp. 1399–1410, December 1995.
- [19] A. Rofougaran, G. Chang, J. J. Rael, J. Y.-C. Chang, M. Rofougaran, P. J. Chang, M. Djafari, J. Min, E. W. Roth, A. A. Abidi, and H. Samueli, “A single-chip 900-MHz spread-spectrum wireless transceiver in 1- μ m CMOS—Part II: Receiver design,” *IEEE J. Solid-State Circuits*, vol. 33, pp. 535–547, April 1998.
- [20] D. K. Shaeffer and T. H. Lee, *Low-Power CMOS Radio Receivers*. Kluwer Academic Publishers, Norwell (MA), 1999.
- [21] W. Thomann, J. Fenk, R. Hagelauer, and R. Weigel, “Fully Integrated W-CDMA IF Receiver and IF Transmitter Including IF Synthesizer and On-Chip VCO for UMTS Mobiles,” *IEEE J. Solid-State Circuits*, vol. 36, pp. 1407–1419, September 2001.
- [22] J. Crols, and M. Steyaert, *CMOS Wireless Transceiver Design*. Kluwer Academic

- Publishers, Dordrecht (The Netherlands), 1997.
- [23] M. S. J. Steyaert, J. Janssens, B. de Muer, M. Borremans, and N. Itoh, "A 2-V CMOS cellular transceiver front-end," *IEEE J. Solid-State Circuits*, vol. 35, pp. 1895–1907, December 2000.
 - [24] S. Tadjpour, E. Cijvat, E. Hegazi and A. A. Abidi, "A 900-Mhz Dual-Conversion Low-IF GSM Receiver in 0.35- μ m CMOS," *IEEE J. Solid-State Circuits*, vol. 36, pp. 1992–2002, December 2001.
 - [25] A. Rofougaran, G. Chang, J. J. Rael, J. Y.-C. Chang, M. Rofougaran, P. J. Chang, M. Djafari, M.-K. Ku, E. W. Roth, A. A. Abidi, and H. Samueli, "A single-chip 900-MHz spread-spectrum wireless transceiver in 1- μ m CMOS—Part I: Architecture and transmitter design," *IEEE J. Solid-State Circuits*, vol. 33, pp. 515–534, April 1998.
 - [26] J. A. Weldon, R. S. Narayanaswami, J. C. Rudell, L. Lin, M. Otsuka, S. Dedieu, L. Tee, K.-C. Tsai, C.-W. Lee, and P. R. Gray, "A 1.75-GHz Highly Integrated Narrow-Band CMOS Transmitter with Harmonic-Rejection Mixers," *IEEE J. of Solid-State Circuits*, vol. 36, pp. 2003–2015, December 2001.
 - [27] F. Behbahani, A. Karimi-Sanjaani, W.-G. Tan, A. Roithmeier, J. C. Leete, K. Hoshino, and A. A. Abidi, "Adaptive Analog IF Signal Processor for a Wide-Band CMOS Wireless Receiver," *IEEE J. Solid-State Circuits*, vol. 36, pp. 1205–1217, August 2001.
 - [28] H. Khorramabadi, M. J. Tarsia, and N. S. Woo, "Baseband Filters for IS-95 CDMA Receiver Applications Featuring Digital Automatic Frequency Tuning," *Digest of Technical Papers, International Solid-State Circuits Conference*, pp. 172-173, February 1996.
 - [29] R. K. Hester, S. Mukherjee, D. Padgett, D. Richardson, W. Bright, M. M. Sarraj, J. T. Nabicht, M. D. Agah, A. Bellaouar, I. Chaudhry, J. R. Hellums, K. Islam, A. Loloee, C.-Y. Tsay, and G. Westphal, "CODEC for Echo-Canceling Full-Rate ADSL Modems," *IEEE J. Solid-State Circuits*, vol. 34, pp. 1973–1985, December 1999.
 - [30] C. Conroy, S. Sheng, A. Feldman, G. Uehara, A. Yeing, C.-J. Hung, V. Subramanian, P. Chiang, P. Lai, X. Si, J. Fan, D. Flynn, and M. He, "A CMOS Analog Front-End IC for DMT ADSL," *Digest of Technical Papers, International Solid-State Circuits Conference*, pp. 240-241, February 1999.
 - [31] H. Darabi, and A. A. Abidi, "A4.5-mW 900-MHz CMOS Receiver for Wireless Paging," *IEEE J. Solid-State Circuits*, vol. 35, pp. 1085–1096, August 2000.
 - [32] T. Hollman, S. Lindfors, M. Länsirinne, J. Jussila, and K. A. I. Halonen, "A 2.7-V CMOS Dual-Mode Baseband Filter for PDC and WCDMA," *IEEE J. Solid-State Circuits*, vol. 36, pp. 1148–1153, July 2001.
 - [33] Z.-Y. Chang, D. Macq, D. Haspeslagh, P. M. P. Spruyt, and B. L. A. G. Goffart, "A CMOS Analog Front-End Circuit for an FDM-Based ADSL System," *IEEE J. Solid-State Circuits*, vol. 30, pp. 1449–1456, December 1995.
 - [34] D. K. Shaaffer, A. R. Shahani, S. S. Mohan, H. Samavati, H. R. Rategh, M. M. Hershenson, M. Xu, C. P. Yue, D. J. Eddleman, and T. H. Lee, "A 115-mW, 0.5- μ m CMOS GPS Receiver with Wide Dynamic-Range Active Filters," *IEEE J.*

- Solid-State Circuits*, vol. 33, pp. 2219–2231, December 1998.
- [35] H. Yamazaki, K. Oishi, and K. Gotoh, “An Accurate Center Frequency Tuning Scheme for 450-kHz CMOS Gm-C Bandpass Filters,” *IEEE J. Solid-State Circuits*, vol. 34, pp. 1691–1697, December 1999.
 - [36] K. Lee, J. Park, J.-W. Lee, S.-W. Lee, H. K. Huh, D.-K. Jeong, and W. Kim, “A Single-Cip 2.4-GHz Direct-Conversion CMOS Receiver for Wireless Local Loop using Multiphase Reduced Frequency Conversion Technique,” *IEEE J. Solid-State Circuits*, vol. 36, pp. 800–809, May 2001.
 - [37] G. Nicollini, P. Confalonieri, C. Crippa, A. Nagari, S. Mariani, A. Calloni, M. Moioli, and C. Dallavalle, “A High-Performance Analog Front-End 14-Bit Codec for 2.7-V Digital Cellular Phones,” *IEEE J. Solid-State Circuits*, vol. 33, pp. 1158–1167, August 1998.
 - [38] C. S. Wong, “A 3-V GSM Baseband Transmitter,” *IEEE J. Solid-State Circuits*, vol. 34, pp. 725–730, May 1999.
 - [39] A. Gattani, D. W. Cline, P. J. Hurst, and P. M. Mosinskis, “A CMOS HDLS2 Analog Front-End,” *IEEE J. Solid-State Circuits*, vol. 35, pp. 1964–1975, December 2000.
 - [40] I. Mehr, P. Maulik, and D. Paterson, “A 12-bit Integrated Analog Front End for Broadband Wieline Networks,” *IEEE J. Solid-State Circuits*, vol. 37, pp. 302–309, March 2002.
 - [41] R. Gregorian, and G. C. Temes, *Analog MOS Integrated Circuits for Signal Processing*. John Wiley, New York, 1986.
 - [42] K. Nishimura, *Optimal Partitioning of Analog and Digital Circuitry in Mixed-Signal Circuits for Signal Processing*, PhD. Dissertation, University of California, Berkeley, 1993.
 - [43] B. Brandt, and B. A. Wooley, “A 50-MHz Multibit $\Sigma\Delta$ Modulator for 12-b 2-MHz A/D Conversion,” *IEEE J. Solid-State Circuits*, vol. 34, pp. 748–760, June 1999.
 - [44] Y. Geerts, A. M. Marques, M. S. J. Steyaert, and W. Sansen, “A 3.3-V, 15-Bit, Delta-Sigma ADC with a Signal Bandwidth of 1.1MHz for ADSL Applications,” *IEEE J. Solid-State Circuits*, vol. 34, pp. 927–936, June 1999.
 - [45] T. Burger, and Q. Huang, “A 13.5-mW 185-Msample/s $\Sigma\Delta$ Modulator for UMTS/GSM Dual-Standard IF Reception,” *IEEE J. Solid-State Circuits*, vol. 36, pp. 1868–1878, December 2001.
 - [46] F. Medeiro, B. Pérez-Verdú, and A. Rodríguez-Vázquez. *Top-down Design of High-Performance Sigma-Delta Modulators*. Kluwer Academic Publishers, Dordrecht (The Netherlands), 1999.
 - [47] E. J. van der Zwan, K. Philips and C. A. A. Bastiaansen, “A 10.7-MHz IF-to-Baselband $\Sigma\Delta$ A/D Conversion System for AM/FM Radio Receivers”, *IEEE J. Solid-State Circuits*, vol. 35, pp. 1810–1819, December 2000.
 - [48] J. van Engelen, R. van de Plassche, E. Stikvoort, and A. Venes, “A Sixth-Order Continuous-Time Bandpass Sigma-Delta Modulator For Digital Radio IF”, *IEEE J. Solid-State Circuits*, vol. 34, pp. 1753–1764, December 1999.
 - [49] L. Breems, and J. H. Huijsing, *Continuous-Time Sigma-Delta Modulation for A/D*

- Conversion in Radio Receivers.* Kluwer Academic Publishers, Dordrecht (The Netherlands), 2001.
- [50] B. Razavi, *RF Microelectronics*, Prentice Hall, Upper Saddle River (NJ), 1998.
 - [51] T. Hanusch, F. Jehring, H.-J. Jentschel, and W. Kluge, "Analog Baseband IC for Dual Mode Direct Conversion Receiver," *Proc. of the European Solid-State Circuits Conference*, pp. 244-246 and 249, September 1996.
 - [52] T.-L. D. Au, *Programmable, Low-noise, High-linearity Baseband Filter for a Fully-Integrated, Multi-Standard, CMOS RF Receiver*. Masters Thesis, University of California at Berkeley, January 1999.
 - [53] H. Elwan, A. M. Soliman, and M. Ismail, "A CMOS Norton Amplifier-Based Digitally Controlled VGA for Low-Power Wireless Applications," *IEEE Trans. Circuits and Systems II*, vol. 48, pp. 245-254, March 2001.
 - [54] J. Jussila, J. Ryyynänen, K. Kivekäs, L. Sumanen, A. Pärssinen, and K. A. I. Halonen, "A 22-mA 3.0dB NF Direct Conversion Receiver for 3G WCDMA," *IEEE J. Solid-State Circuits*, vol. 36, pp. 2025-2029, December 2001.
 - [55] J. Rudell, J. Weldon, J. Ou, L. Lin, and P. R. Gray, "An integrated GSM/DECT receiver: Design specification," UCB Electronics Laboratory Memo., Berkeley, CA, 1998.
 - [56] S. Fitz, "Receiver architectures for GSM handsets," in *IEE Coll. Design of Digital Cellular Handsets*, London, U.K., 1998, pp. 3/1-3/10.
 - [57] W. H. Sansen, and R. G. Meyer, "Distortion in bipolar transistor variable gain amplifiers," *IEEE J. Solid-State Circuits*, vol. 8, pp. 275-282, August 1973.
 - [58] F. Behbahani, Y. Kishigami, J. Leete, and A. A. Abidi, "CMOS 10 MHz-IF down-converter with on-chip broadband circuit for large image-suppression," in *Proc. Symp. VLSI Circuits*, Kyoto, Japan, 1999, pp. 83-86.
 - [59] I. Mehr, P. C. Maulik, and D. Paterson, "A 12-bit Integrated Analog Front End for Broadband Wireline Networks," *IEEE J. Solid-State Circuits*, vol. 37, pp. 302-309, March 2002.
 - [60] K. R. Laker and W. M. C. Sansen, *Design of Analog Integrated Circuits and Systems*. McGraw-Hill, New York, 1994.
 - [61] R. Schaumann, M. S. Ghausi and K. R. Laker, *Design of Analog Filters*. Prentice-Hall, Englewood Cliffs, 1990.
 - [62] K. D. Peterson, A. Nedungadi and R. L. Geiger, "Amplifier Design Considerations for High Frequency Monolithic Filters," *Proceedings of the 1987 European Conf. on Circuit Theory and Design*, pp. 321-326, September 1987.
 - [63] D. Leenaerst, J. van der Tang, and C. Vaucher, *Circuit Design for RF Transceiver*. Kluwer Academic Publishers, Dordrecht (The Netherlands), 2001.
 - [64] J. Janssens and M. Steyaert, *CMOS Cellular Receiver Front-Ends: From Specification to Realization*. Kluwer Academic Publishers, Dordrecht (The Netherlands), 2002.

Chapter 14

Design Methodologies for Sigma-Delta Converters

*Francisco V. Fernández, Rocío del Río, Rafael Castro-López,
Oscar Guerra, Fernando Medeiro and Belén Pérez-Verdú*

15.1 INTRODUCTION^{†1}

Oversampling converters have become very popular due to their ability to solve problems found in other architectures, like the need for high-accuracy analog anti-aliasing filtering and the large sensitivity to circuit imperfections and noisy environments.

The use of a high sampling frequency (compared to the signal bandwidth) relaxes the specifications of the antialiasing filter, so that a first-order analog filter is usually sufficient and the heaviest filtering tasks are shifted to the digital domain. Furthermore, by combining oversampling and sigma-delta modulation, it is possible to design data converters with large resolution, robust operation and relative insensitivity to non-idealities. This makes oversampling converters, and especially those that incorporate $\Sigma\Delta$ modulation, very appropriate for integration in standard digital CMOS processes, where very fast and dense digital circuits can be implemented, whereas accurate analog performance is not easily obtained.

However, the large conceptual and practical differences between traditional Nyquist-rate converters and those based on $\Sigma\Delta$ modulation render the reuse of the design methodologies developed for the former ones impossible [1],[2].

Although it is generally accepted that $\Sigma\Delta$ converters are intrinsically less sensitive to the physical circuit non-idealities than other types of data converters, it is also true that the error mechanisms associated to such non-idealities must be taken

1. **Acknowledgement:** This work has been partially funded by the Spanish Ministry of Science and Technology (with support from the European Regional Development Fund) under contract TIC2001-0929 and the Commission of the EU in the framework of the Projects ESPRIT IV 29648 RAPID and IST 2001-34283 TAMES-2.

into account for electrical implementations. Furthermore, the robustness or insensitivity concepts may become quite fallacious when dealing with high-performance $\Sigma\Delta$ converters. In conclusion, practical design methodologies for high-performance $\Sigma\Delta$ converters require a deep knowledge of the modulator operation and how this is affected by the circuit non-idealities.

This chapter focuses on switched-capacitor $\Sigma\Delta$ converters, which has been the technique used in most converter implementations. However, many concepts and ideas are also applicable to switched-current and continuous-time implementations. Section 15.2 in this chapter introduces a modern top-down design methodology applied to the design of $\Sigma\Delta$ converters. The most relevant tasks of this design methodology, namely simulation, synthesis and layout generation, are discussed in sections 15.3 to 15.5.

15.2 TOP-DOWN DESIGN METHODOLOGY

Analog and mixed-signal CAD tools are far behind their digital counterparts. Digital design benefits from a well-established hierarchically-structured design. Various abstraction levels are available for the different hierarchical levels and a set of standard rules enables the transmission of specifications through the hierarchy. Moreover, if these rules are observed, the correct functionality of the lower level blocks ensures that of the upper level blocks. Such a rigorous and structured approach does not exist for analog and mixed-signal systems. Hierarchy in these systems obeys mainly to a structural decomposition, signal abstraction levels do not exist, and neither do sets of clear rules to ensure proper operation through the hierarchy.

In spite of these difficulties, it is already commonly accepted that systematic design of complex mixed-signal systems can only be realistically addressed by means of a hierarchical design methodology. This design methodology proceeds through a top-down synthesis flow across the different hierarchical levels. At each level, an appropriate architecture is selected for a given set of block specs. Such block specs are then transmitted to specs of the component sub-blocks. Block performance is verified at this level and, if correct, the synthesis process continues for the component sub-blocks at the lower hierarchical level. The top-down process ends at the device level and, then, a bottom-up layout generation and verification flow starts. At each level, the component blocks are laid out according to the selected architecture, checked, extracted and verified using appropriate simulation tools. Multiple re-design iterations are performed at the different levels according to the verification results.

Fig. 15.1 shows the adaptation of this top-down methodology to the design of

$\Sigma\Delta$ converters. Starting from the converter specifications, i.e., speed, resolution, etc., an appropriate topology is selected. Let us assume that this corresponds to the cascade modulator shown in Fig. 15.2. This architecture is going to be used as a representative example in different sections of this chapter. It is composed of integrators, comparators, multibit quantizers, D/A converters and digital blocks. The converter specifications are transmitted to specifications for these blocks. For each of them, e.g., the integrator, specification transmission corresponds to obtaining switch and capacitor sizes and amplifier specs for the selected architecture. Before proceeding to the following hierarchical level, performance of the converter is verified for the transmitted specifications. For each sub-block, i.e., the amplifier, an appropriate architecture is selected and specification transmission at this lowest hierarchical level directly yields device sizes and bias voltages/currents. Then, layout phase is started. For the considered example, the opamp is laid out, extracted and electrically simulated and compared with the block specifications. Resulting layout is joined to switch and capacitor layout, yielding the integrator layout, which is again extracted and verified. Then, layout is joined to layouts of other integrators, comparators, D/A converters, etc.

Practical implementation of this methodology can take into account multiple re-design iterations in case of non-fulfillment of performance specs (either in validation steps during the top-down synthesis process or during the bottom-up layout phase), as well as tight interactions between synthesis and layout phases, even between different hierarchical levels.

15.3 SIMULATION

Accurate and efficient simulation is a crucial step in any design methodology for electronic circuits. Simulation is an irreplaceable tool for verification of circuit performances before they are fabricated. But it is not limited to this task, simulation can also be used for design space exploration: by performing repetitive executions, parametric sweeps, etc. Finally, it is also at the core of many current synthesis and optimization strategies which include some evaluation of circuit performances within an iterative loop.

Traditionally, electrical simulators have been the most widely used simulation tools. Circuits are built of an interconnection of basic devices, for each of which a model is provided, and the complete circuit is simulated using a SPICE-like simulator. Simulation accuracy is high, provided that accurate device models are used, but computation time grows with circuit size and the type and length of simulation. There is no doubt that electrical simulation is the logical choice for simulating basic cells. A wide variety of commercial tools exists: HSPICE, SPECTRE,

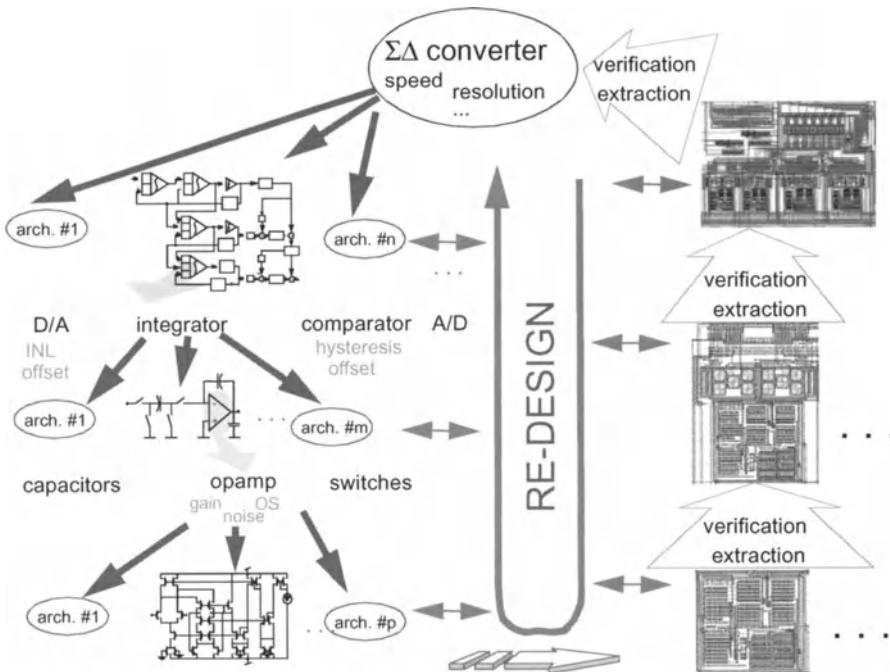


Figure 15.1: Top-down design methodology for $\Sigma\Delta$ converters.

PSPICE, ELDO, etc.

However, the situation changes considerably when dealing with higher-level sub-systems, like data converters. Electrical simulation of Nyquist-rate converters is not practical due to the large circuit size: thousands of transistors in typical cases. Unlike Nyquist-rate converters, circuitry in $\Sigma\Delta$ converters is not very complex – typically a few hundredths of transistors. Therefore, circuit size seems not to be an obstacle for the electrical level simulation of these circuits. However, electrical simulation of $\Sigma\Delta$ converters is not practical either, although due to very different reasons. $\Sigma\Delta$ modulators contain strongly non-linear blocks, i.e., the quantizer. Therefore, small-signal analysis is not possible and transient simulation becomes the only choice. Extraction of converter performances requires simulation over several cycles of the input signal. However, due to the oversampling nature of the system, there are a rapidly varying clock and a slowly varying input signal. Therefore, transient simulation over tens or hundredths of thousands of clock cycles must be performed. The consequence is that a single simulation of a $\Sigma\Delta$ modulator at the electrical level can take days or weeks of CPU time [3].

A conventional solution to avoid too lengthy electrical simulations of analog circuits is to use macromodels of the basic cells. Although the problem is simplified, computation times are still too long because the equation formulation is still a set of differential equations which must be solved by numerical integration using millions of time-steps. A more evolved approach is to resort to mixed-mode simulators. Critical parts are modeled at the device level while behavioral models are used for the rest of the system. However, the numerical solution of the analysis equations is still costly in CPU time.

The last and most efficient solution is event-driven behavioral level simulation, which brings the simulation of $\Sigma\Delta$ modulators closer to logic simulation. This approach requires that the circuit can be partitioned into basic blocks with independent functionality. This implies that an instantaneous block output cannot be related to itself, that is, either there is no global feedback loop, or, in case such a loop exists, there is a delay that avoids the instantaneous dependence. The behavioral simulation of such circuits requires a behavioral model for each block, in the form of explicit expressions relating the output variables with the input and internal state variables.

Several tools following this approach have been reported [1],[4]-[7]. They differ mainly in the topologies/basic blocks included, the friendliness of the user interface, the accuracy of the behavioral models and the postprocessing capabilities.

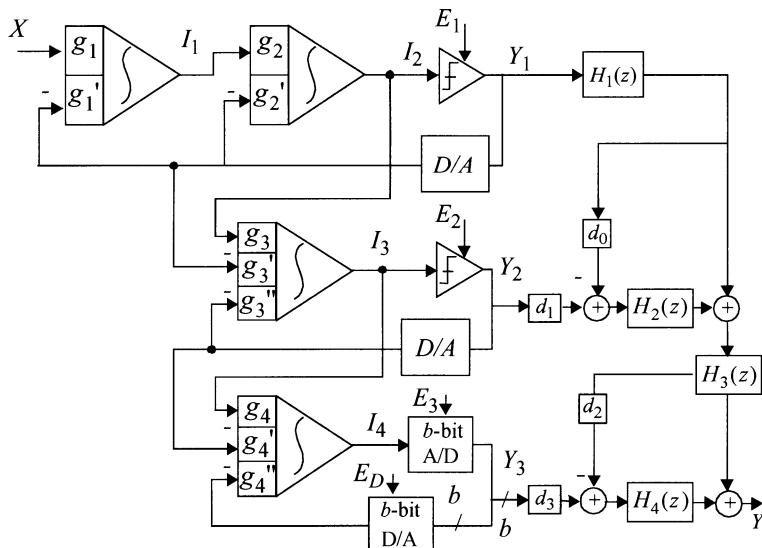


Figure 15.2: Fourth-order three-stage multi-bit $\Sigma\Delta$ modulator (2-1-1mb).

An interesting alternative is the look-up table method [8]. It simulates the performance of subcircuits for a variety of inputs and initial conditions using an electrical simulator, and the results are stored in a table. During the execution of the behavioral simulation, the output of such subcircuits are computed by interpolation of the look-up table data. Since the generation of tables for each block is relatively expensive and a new table has to be generated whenever the architecture itself or any circuit parameter is changed, this approach is very attractive for final verification but not for earlier stages of the converter design. In this sense, it becomes a complementary strategy to conventional event-driven simulation approaches.

The behavioral simulator ASIDES [1],[6] is going to be used as a reference for further description. ASIDES contains models for a wide set of sub-block non-idealities, it is able to use look-up table models for any sub-block and it has a wide range of postprocessing capabilities.

15.3.1 Analytical models for event-driven behavioral simulation

The fundamental blocks of SC $\Sigma\Delta$ modulators are integrators, single-bit or multi-bit quantizers, D/A converters and digital blocks. Through the appropriate interconnection of these few blocks a large number of architectures can be obtained, whose behavior can be affected by the non-idealities of each block. Assuming that the non-ideal behavior of the digital circuitry does not degrade the performance of the modulator we will focus on the circuitry with analog part.

15.3.1.1 Integrator model

The integrator can be considered as the fundamental block because, due to its position, its non-idealities directly affect the performance of $\Sigma\Delta$ modulators.

Transient response

Defective settling is becoming one of the dominant limiting factors in SC $\Sigma\Delta$ modulators, especially for those employing high sampling frequencies, i.e., high-speed high-resolution modulators required by digital subscriber line specifications. Several models have been reported which take into account the amplifier bandwidth and slew-rate limitations during the integration phase and the incomplete settling of the voltage in the input capacitor due to the non-zero ON resistance of the analog switches in the sampling phase, neglecting, therefore, the finite sampling dynamics. This may yield important underestimates of defective settling, especially in high speed $\Sigma\Delta$ modulators [9],[10].

To model the dynamic behavior of the integrator let us consider the generic

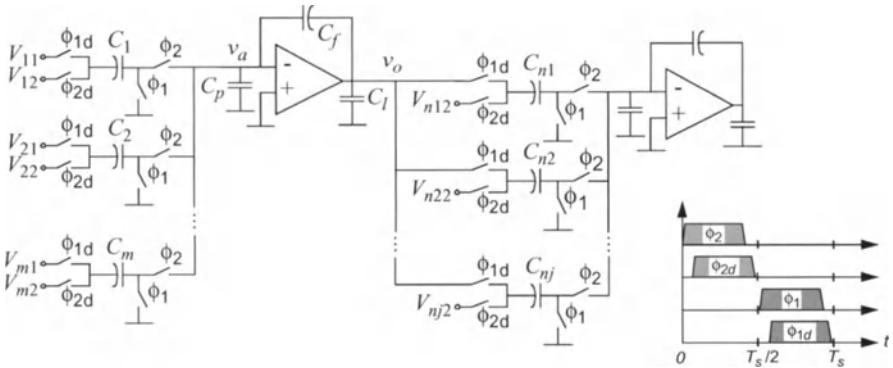


Figure 15.3: SC integrator model.

model in Fig. 15.3. It contains:

- m input branches connected to switching levels V_{k1} and V_{k2} ,
- the parasitic capacitor C_p associated to the summation node,
- the capacitive load C_l associated to the amplifier output node and to the bottom plate of the integration capacitor C_f , and
- j branches of an integrator loading its output during sampling, and switching to levels V_{nk2} during integration.

ϕ_{1d} and ϕ_{2d} in Fig. 15.3 represent two slightly delayed versions of the clock phases ϕ_1 and ϕ_2 , used to prevent signal-dependent charge injection [11]. From the modeling point of view, the shorter time available for integrator settling must be taken into account.

On the other hand, the amplifier is modeled with a single-pole dynamic, and a non-linear characteristic, with maximum output current I_o , as Fig. 15.4 shows.

Let us denote $v_{a,n-1}$ and $v_{o,n-1}$ the opamp input and output voltages at the end of the $(n-1)$ -th sampling phase. Charge conservation at the beginning of the integration phase, $t = 0$, determines a jump on these voltages, in the opposite direction to their final values, to:

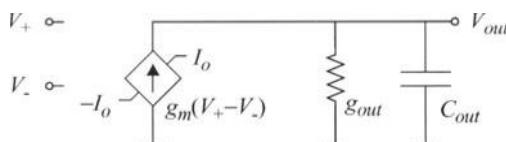


Figure 15.4: Opamp model.

$$\begin{aligned} v_{ai,i} &= \frac{1}{C_{eq,i}} \left(1 + \frac{C_l}{C_f} \right) \sum_{k=1}^m (V_{k2} - V_{k1}) C_k + \frac{C_p + C_l(1 + C_p/C_f)}{C_{eq,i}} v_{a,n-1} \\ v_{oi,i} &= v_{o,n-1} + (C_f/(C_f + C_l))(v_{ai,i} - v_{a,n-1}) \end{aligned} \quad (15.1)$$

where $C_{eq,i}$ is the equivalent capacitive load at the opamp output during integration, given by:

$$C_{eq,i} = C_p + \sum_{k=1}^m C_k + C_l \left[1 + \left(C_p + \sum_{k=1}^m C_k \right) / C_f \right] \quad (15.2)$$

Notice that $v_{a,n-1} \neq 0$ to reflect a possible defective settling by the end of the preceding sampling phase. Depending on $v_{ai,i}$, two possibilities can be identified:

(a) $|v_{ai,i}| \leq I_o/g_m$, with the amplifier operating linearly:

$$v_a(t) = v_{ai,i} \exp(-g_m/C_{eq,i}t) \quad (15.3)$$

(b) $|v_{ai,i}| > I_o/g_m$, so that the amplifier slews with constant slope:

$$v_a(t) = v_{ai,i} - \frac{I_o}{C_{eq,i}} \operatorname{sgn}(v_{ai,i}) t \quad (15.4)$$

The slewing continues until $t = t_{o,i}$, when the condition for the amplifier to start operating linearly, $v_a(t_{o,i}) = I_o/g_m$, fulfills

$$t_{o,i} = \frac{C_{eq,i}}{I_o} |v_{ai,i}| - \frac{C_{eq,i}}{g_m} \quad (15.5)$$

and from then on $v_a(t)$ will relax exponentially:

$$v_a(t) = \frac{I_o}{g_m} \operatorname{sgn}(v_{ai,i}) \exp\left[-\frac{g_m}{C_{eq,i}}(t - t_{o,i})\right] \quad (15.6)$$

During the integration phase $v_o(t)$ is given by

$$v_o(t) = v_{o,n-1} - \left(1 + \frac{C_p}{C_f} \right) v_{a,n-1} - \sum_{k=1}^m \frac{C_k}{C_f} (V_{k2} - V_{k1}) + \left[1 + \frac{C_p + \sum_{k=1}^m C_k}{C_f} \right] v_a(t) \quad (15.7)$$

where $v_a(t)$ stands for eqs. (15.3) and (15.6) or (15.4) depending, respectively, on the amplifier linear operation, partial or complete slewing.

Let us denote $v_a(T_s/2)$ and $v_o(T_s/2)$ the opamp input and output voltages at the end of the integration phase. Charge conservation at the beginning of the sampling phase, $t = T_s/2$, determines a new voltage jump to

$$v_{ai,s} = v_a(T_s/2) - \sum_{k=1}^j \left[\frac{C_{nk}}{C_{eq,s}} (v_o(T_s/2) - V_{nk2}) \right] \quad (15.8)$$

$$v_{oi,s} = v_o(T_s/2) + (1 + C_p/C_f)[v_{ai,s} - v_a(T_s/2)]$$

where $C_{eq,s}$ is the equivalent capacitive load during sampling:

$$C_{eq,s} = C_p + \left(C_l + \sum_{k=1}^j C_{nk} \right) \left(1 + \frac{C_p}{C_f} \right) \quad (15.9)$$

Operating in a similar way as for the integration phase, we get:

$$v_a(t) = v_{ai,s} \exp \left[-\frac{g_m}{C_{eq,s}} \left(t - \frac{T_s}{2} \right) \right] \quad \text{if } |v_{ai,s}| \leq I_o/g_m \quad (15.10)$$

$$v_a(t) = v_{ai,s} - \frac{I_o}{C_{eq,s}} \operatorname{sgn}(v_{ai,s}) \left(t - \frac{T_s}{2} \right) \quad \text{if } \begin{cases} |v_{ai,s}| > I_o/g_m \\ t < t_{o,s} \end{cases} \quad (15.11)$$

$$v_a(t) = \frac{I_o}{g_m} \operatorname{sgn}(v_{ai,s}) \exp \left[-\frac{g_m}{C_{eq,s}} (t - t_{o,s}) \right] \quad \text{if } \begin{cases} |v_{ai,s}| > I_o/g_m \\ t \geq t_{o,s} \end{cases} \quad (15.12)$$

where

$$t_{o,s} = \frac{T_s}{2} + \frac{C_{eq,s}}{I_o} |v_{ai,s}| - \frac{C_{eq,s}}{g_m} \quad (15.13)$$

During the sampling phase $v_o(t)$ is given by

$$v_o(t) = v_o(T_s/2) + (1 + C_p/C_f)[v_a(t) - v_a(T_s/2)] \quad (15.14)$$

where $v_a(t)$ stands for eqs. (15.10) and (15.12) or (15.11) depending on the amplifier linear operation, partial or complete slewing.

The single-pole model has been widely used and usually suffices because single-stage amplifiers with a single dominant pole are commonly used. However, two-pole models are convenient even for single-stage amplifiers with small phase margin, which often occurs in high frequency applications in which the dynamic

properties of the transistors must be fully exploited [1].

Circuit noise

By using event-driven behavioral simulation, only those phenomena that present changes with a frequency smaller or equal to that of the clock can be modeled. Furthermore, the modeling of continuous nature mechanisms is possible only when their influence is circumscribed to each individual clock cycle, or in other words, such influence can be synchronously partitioned. This is the case, for example, of the integrator settling described above.

This does not occur with circuit noise: first, it is a phenomenon of continuous nature that varies randomly and second, its cut-off frequency is generally well above the sampling frequency. The first problem is partially solvable because, though computers base their operation on deterministic algorithms, it is possible to build routines that generate random numbers with very low self-correlation. So, the principal difficulty is the timing: the reduction of the internal time-step of a computation algorithm in order to adapt it to the changing speed of the signals it handles is used in numerical integration algorithms. Thus, the calculation speed – the main advantage of the behavioral simulation – is definitively lost.

However, since the clock frequency in oversampling converters is several times larger than that of the signals being processed, the white noise modulation leads to an input-equivalent noise that, at least in the frequencies of interest, presents a constant spectral density. The modulation of this noise is effective only for frequencies near the sampling frequency. By taking advantage of this fact, the modulator input-equivalent circuit noise can be pre-calculated and injected at the clock rate using a random number generator. Thus, a power spectral density is obtained that equals that of the modulated circuit noise in the low-frequency region.

Let us consider the two-branch SC integrator in Fig. 15.5^{†1}, where we assume that the signal sampling capacitor, C_1 , can be different from the feedback sampling capacitor, C_f in order to create an integrator input-out gain $\xi = C_1/C_f$ different from one^{†2}. In this circuit, the noise contributors are the OTA (white + flicker noise) and the switch on-resistance (thermal noise). Whereas flicker noise can be very efficiently removed from the signal band by applying ingenious chopper techniques [12] for example, the low-frequency white noise power spectral

1. For the sake of simplicity, only two branches are considered. The analysis that follows can be easily extended to an arbitrary number of branches.
2. This is an strategy to boost the converter input signal when it comes from sensing devices, for example.

density (PSD) is boosted by the well-known aliasing mechanism: due to dynamic requirements, the equivalent noise bandwidths for these contributions are well in excess of half the sampling frequency, thus provoking undersampling and subsequent aliasing of the noise PSD. A careful analysis shows that the input-equivalent white noise PSD per fully-differential branch can be approximated (at low frequencies) by,

$$S_{in, C_j}(f) \approx \frac{4kT}{C_j f_s} + \frac{4kT(1+n_t)}{3C_{eq, j} f_s} \quad j = 1, 2 \quad (15.15)$$

where k is the Boltzmann constant and T is the absolute temperature. The first term corresponds to the contribution of all the switches in each branch, the second one is for the OTA contribution, where

$$C_{eq, i} = C_p + C_2(1+\xi) + C_f \left[1 + \frac{C_p + C_2(1+\xi)}{C_f} \right] \quad (15.16)$$

is the OTA equivalent load during the integration phase (ϕ_2 high). In estimating the white noise PSD of the OTA, the contributions of MOS devices other than the input ones are compiled in the factor n_t , which equals the summation of the respective transconductance ratios (smaller than unity).

By neglecting the integrator noise contributions other than the first one, the output white noise power of the $\Sigma\Delta$ modulator is obtained:

$$P_{wn} \approx \frac{4kT}{MC_{12}}(1+\xi) + \frac{4kT(1+n_t)}{3MC_{eq, i}}(1+\xi)^2 \quad (15.17)$$

where M stands for the oversampling ratio. In obtaining (15.17), the OTA noise contributions in (15.15) have been considered fully correlated from branch to

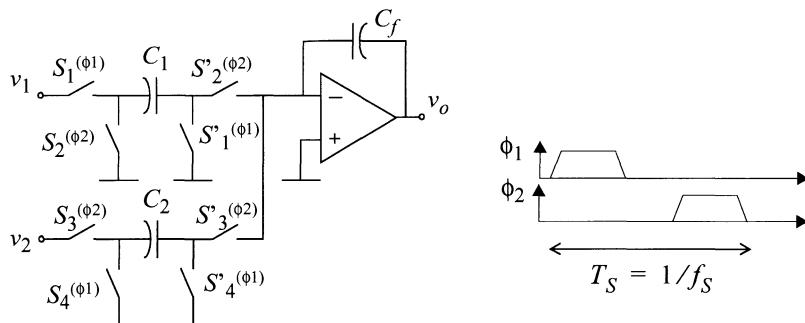


Figure 15.5: Two-branch SC integrator and clock phases.

branch because it is the same circuitry and its noise is sampled at the same instant.

Eq. (15.15) can be used for injecting the white noise in each branch during the event-driven simulation, making use of the well-known relationship for the PSD of a series of random numbers uniformly distributed in the interval $[-\Delta/2, \Delta/2]$ sampled at f_s , i.e., $S_n(f) = \Delta^2/(12f_s)$ [13]. In order to inject the switch contribution – first term in (15.15) – a random number must be generated per branch. However, to reflect the above mentioned correlation, the random number generated for injecting the OTA contribution must be the same for all branches.

Finite and non-linear opamp open-loop DC gain

Ideally, the finite-difference equation that represents the behavior of the SC integrator in Fig. 15.3 is

$$v_{o,n}C_f = \sum_{k=1}^m C_k(v_{k1,n-1} - v_{k2,n}) + v_{o,n-1}C_f \quad (15.18)$$

Considering a finite gain model for the amplifier, the operation of the integrator can be approximated by

$$v_{o,n} \cong \frac{A_V \sum_{k=1}^m g_k(v_{k1,n-1} - v_{k2,n})}{A_V + 1 + \sum_{k=1}^m g_k} + \frac{(A_V + 1)v_{o,n-1}}{A_V + 1 + \sum_{k=1}^m g_k} \quad g_k = \frac{C_k}{C_f} \quad (15.19)$$

Not only the open-loop gain of the amplifiers is finite but it also varies with the DC output voltage. Assuming a polynomial dependence of the gain of the amplifier with the output voltage as

$$A_V = A_0(1 + \gamma_1 v_{o,n} + \gamma_2 v_{o,n}^2 + \dots) \quad (15.20)$$

sequential application of (15.19) and (15.20) leads quickly to the final value of $v_{o,n}$ after a few iterations.

Capacitor non-linearity

Three causes of distortion exist: slew-rate in the settling of the output voltage, non-linear open-loop gain of the amplifier and capacitor non-linearities. The first two mechanisms have already been introduced above. Let us consider again the SC integrator in Fig. 15.2 to model the capacitor non-linearity.

If the capacitors are non-linear, their capacitance varies with the stored voltage:

$$C(v) = C^o(1 + \alpha v + \beta v^2 + \dots) \quad (15.21)$$

A model with a first and a second-order nonlinearity is accurate enough in practice and (15.18) becomes:

$$v_{o,n} = \left[\sum_{k=1}^m v_{k1,n-1} C_k^o \left(1 + \frac{\alpha}{2} v_{k1,n-1} + \frac{\beta}{3} v_{k1,n-1}^2 \right) - \sum_{k=1}^m v_{k2,n} C_k^o \left(1 + \frac{\alpha}{2} v_{k2,n} + \frac{\beta}{3} v_{k2,n}^2 \right) + v_{o,n-1} C_f^o \left(1 + \frac{\alpha}{2} v_{o,n-1} + \frac{\beta}{3} v_{o,n-1}^2 \right) \right] / \left[C_f^o \left(1 + \frac{\alpha}{2} v_{o,n} + \frac{\beta}{3} v_{o,n}^2 \right) \right] \quad (15.22)$$

This equation must be solved numerically: however, for weak non-linearity, as observed in practice, where $\alpha \ll 1$ and $\beta \ll 1$, a simple fast-convergence iterative procedure permits the calculation of the final value of $v_{o,n}$.

The computation of the effects of the capacitor and opamp open-loop gain nonlinearities are included simultaneously in the flow graph of Fig. 15.6. After initializing the input voltages, the integrator internal state and the amplifier gain to their nominal values, the value of $v_{o,n-1}$, $v_{j1,n-1}$ and $v_{j2,n}$ is modified according to the current value of A_V . Then, the value of $v_{o,n}$ is calculated using (15.22), supposing in addition that the denominator of this expression does not depend on $v_{o,n}$. Subsequently $v_{o,n}$ is updated for the actual values of α and β and, finally, the new value of the open-loop gain is calculated via (15.20). This procedure usually converges after two or three iterations which enables an efficient computation.

Jitter noise

In practice, the sampling period is not constant due to intrinsic uncertainties in the time instant in which clock transitions occur. The result is a non-uniform sampling, responsible for the degradation of the quantization noise-shaping function. The larger the frequency of the input signal, the larger such degradation is. Although this error mechanism is due to the integrator, a convenient behavioral model is to consider it as an integral part of the sinusoidal signal generator: a modulator with jitter excited by a pure sinusoidal signal is equivalent to a simulator without jitter excited by a contaminated sinusoidal signal. This equivalence implies that only the contribution of the first integrator is taken into account. The modulator input voltage provided by the sinusoidal generator with jitter becomes:

$$v_i(nT_S) = A \sin[2\pi(nT_S + \Delta t)] \quad (15.23)$$

where Δt has a Gaussian distribution with zero mean and standard deviation σ_t :

$$\Delta t = gauss(0, \sigma_t) \quad (15.24)$$

The complete integrator model containing all non-idealities is computed according to the flow-graph in Fig. 15.7.

15.3.1.2 Quantizer and D/A converter models

In the single-bit case, the quantizer reduces to a comparator. The single-bit D/A

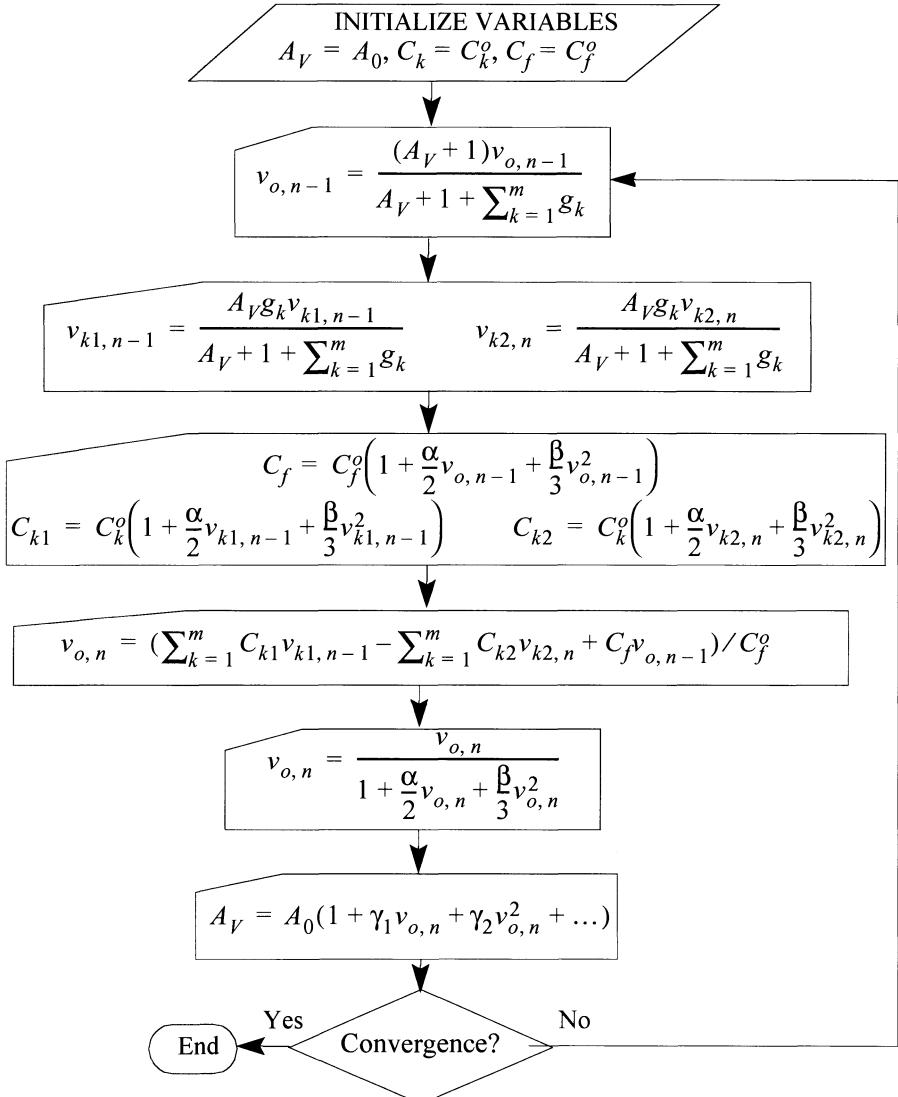


Figure 15.6: Calculation of the integrator output voltage in presence of non-linearities.

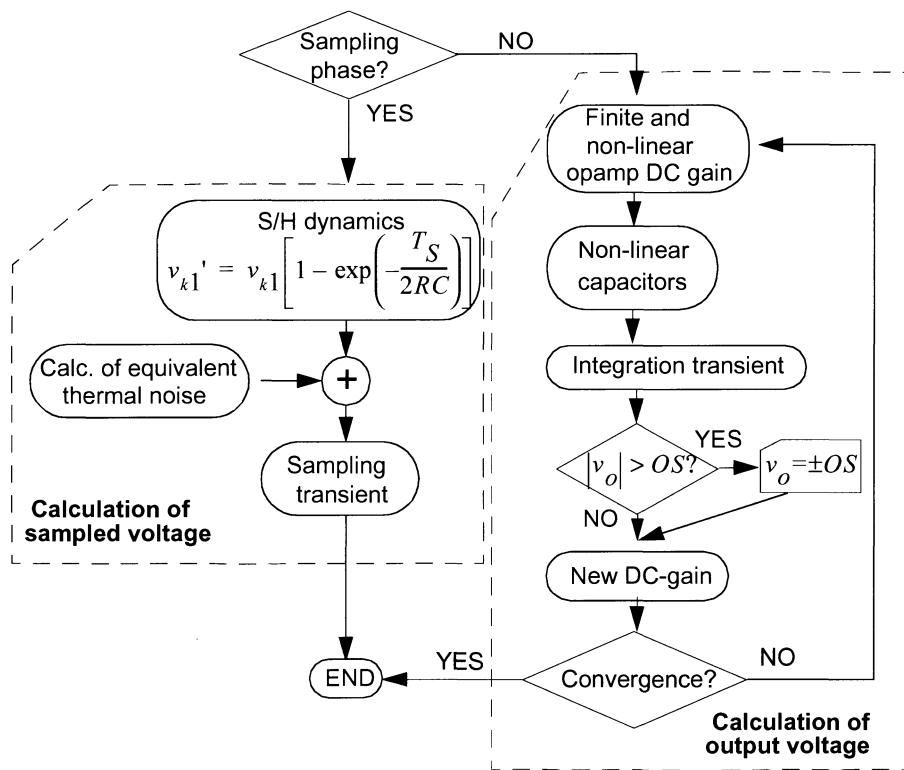


Figure 15.7: Flow diagram of the integrator model.

converter is a very simple and linear block which does not introduce any non-linearity error. The most important non-idealities in the comparator are input offset and hysteresis. The hysteresis arises because the comparator has a memory of the previous state, an overdrive being necessary to make it commute to the correct state. Both, offset and hysteresis, are graphically displayed in Fig. 15.8. In addition to this deterministic hysteresis, another one, of random nature, appears in latched comparators. Due to the non-idealities of real devices, the memory of the previous state is not completely eliminated during the reset phase of this type of comparators. The output of the comparator is determined not only by its input but also by the previous state and the transitions of the signals resetting the latch, leading to an uncertainty zone in the transition between both states. All these non-idealities are covered in the flow graph in Fig. 15.9. It must be noticed that both, offset and hysteresis, are attenuated by the DC gain(s) of the integrator(s) that precede it in the loop and, therefore, they are much less important than the non-ideal-

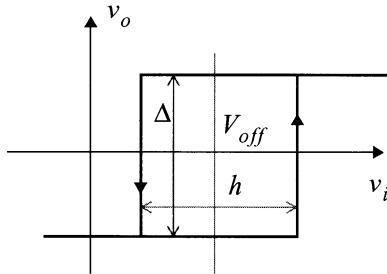


Figure 15.8: Transfer curve of a comparator with hysteresis and offset.

ties of the first integrator.

In the multibit case, offset, gain error and non-linearity are the major non-idealities in the A/D and D/A conversion of the signals in the feedback loop. Among them, the non-linearity of the D/A converter is the most important one, since in many architectures the D/A conversion error is directly added to the modulator input and it appears at the output as distortion. Novel techniques and architectures, such as dynamic matching techniques and dual quantization architectures, partially correct these distortion problems. From the point of view of the behavioral simulation, the non-idealities of multibit A/D and D/A converters are modeled in ASIDES as shown in Fig. 15.10. For a D/A converter, characterized by an offset off , a gain γ and an integral non-linearity INL, the converter input is passed

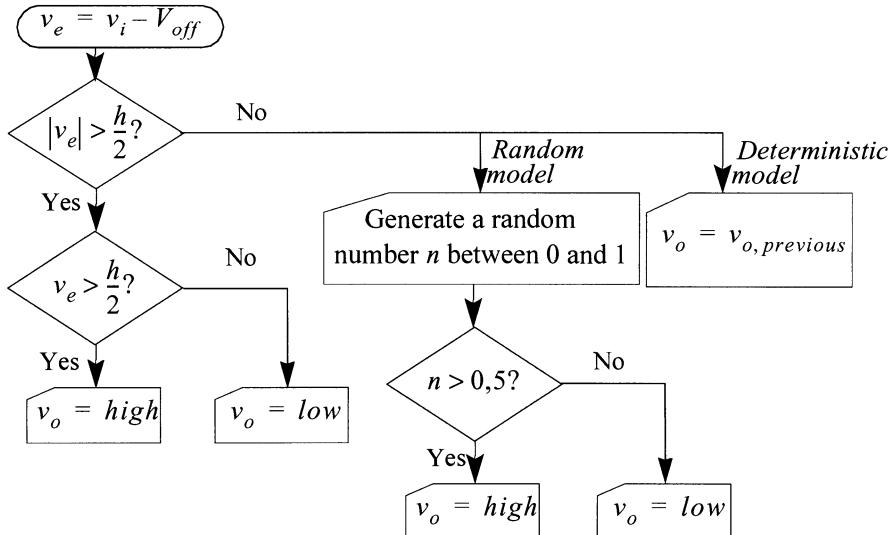


Figure 15.9: Comparator model flow graph.

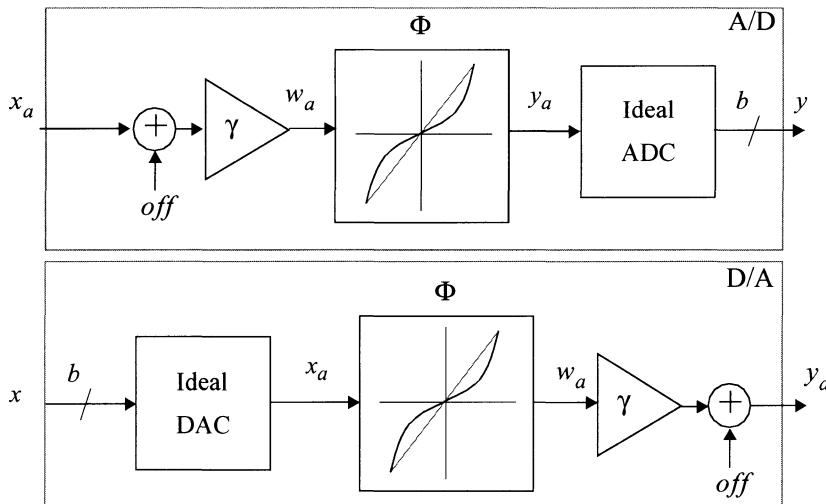


Figure 15.10: Model for the behavioral simulation of multibit A/D and D/A converters.

through an ideal D/A converter. The result goes through a non-linear block with a third-order non-linearity and a gain block. Finally, the offset error is added. An analogous scheme is used for the A/D converter.

15.3.2 Look-up tables

In the previous section, circuit non-idealities have been modeled analytically. For instance, a polynomial model was used for the non-linear open-loop gain of the amplifiers. Fig. 15.11 shows electrical simulation examples of the gain of an operational amplifier versus the DC output voltage for two different process and temperature conditions. The polynomial dependence in section 15.3.1.1 is only a good approximation for small voltage excursions around the central point. But in low-voltage implementations small-gain regions of the DC curve are often visited during the normal modulator operation, so that simple polynomial models are not able to fully predict the impact of the non-linearity. Higher accuracy is obtained by resorting to a table look-up procedure from opamp DC curves obtained by electrical simulation.

Defective settling is another major error source in high-speed modulators. In this case, the analytical model in section 15.3.1 is appropriate for design space exploration and synthesis tasks but for accurate performance verification a table look-up method is also convenient. A significant difference with the non-linear

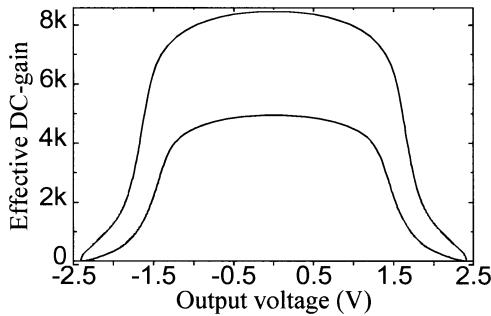


Figure 15.11: DC-gain non-linearity vs. output voltage.

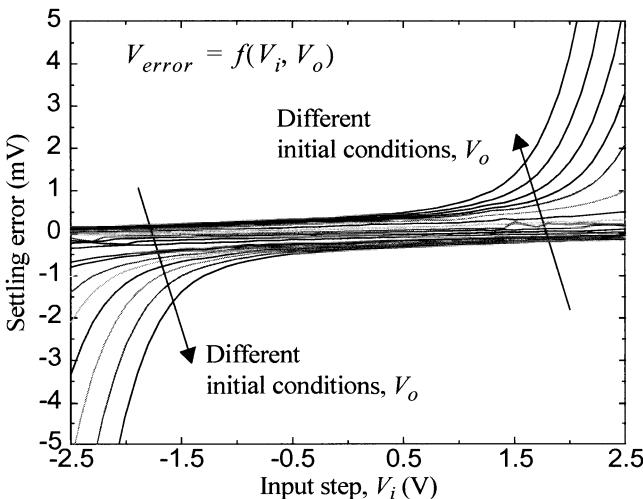


Figure 15.12: Integrator settling error vs. input voltage.

gain of the amplifiers is that the integrator output not only depends on the input voltage but also on the initial conditions, as Fig. 15.12 shows. This means that bi-dimensional tables must be built.

15.3.3 Postprocessing

The outcome of the behavioral simulation is a time-domain series that is digitally processed to perform different types of analysis, as schematically shown in Fig. 15.13. Dynamic analysis includes output spectrum, SNDR as a function of the input level, frequency, etc., in-band error power, effective resolution. Static analysis allows the evaluation of the DC transfer curve of the converter. Monte

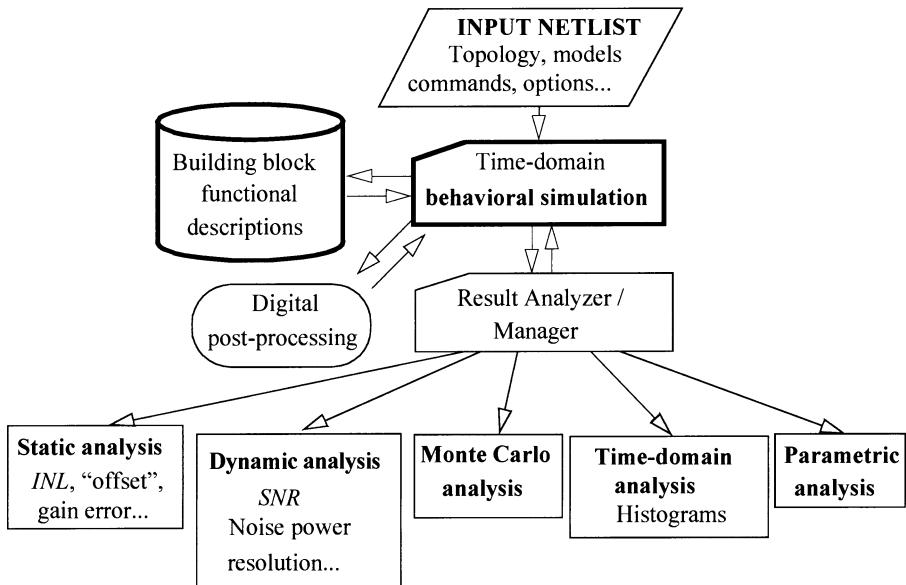


Figure 15.13: ASIDES architecture.

Carlo analysis allows evaluating the impact of fluctuations in integrator gains or basic block specifications on the modulator performance. Finally, parametric analysis allows performing previous analyses with some non-idealities defined as variable parameters. Unlike previous types of analysis, parametric analysis is not intended for performance verification but for design space exploration, including topology selection.

15.3.4 Implementation and examples

Event-driven $\Sigma\Delta$ behavioral simulation has been usually implemented as special-purpose simulators. In them, both, simulation engine and models, are implemented using a common programming language like C. They have a large flexibility to describe circuit non-idealities and are very fast, which make them appropriate even for synthesis tasks based on simulation-in-a-loop strategies [1]. The major drawback is that they are restricted to this kind of systems and are hardly integrated with other simulation tools.

As a simulation example, consider the fourth-order cascade $\Sigma\Delta$ modulator in Fig. 15.2. Fig. 15.14 illustrates the analysis and postprocessing capabilities of ASIDES by showing the effect of integrator weight mismatch by means of the Monte Carlo simulation of the signal-to-noise+distortion ratio (SNDR) as a func-

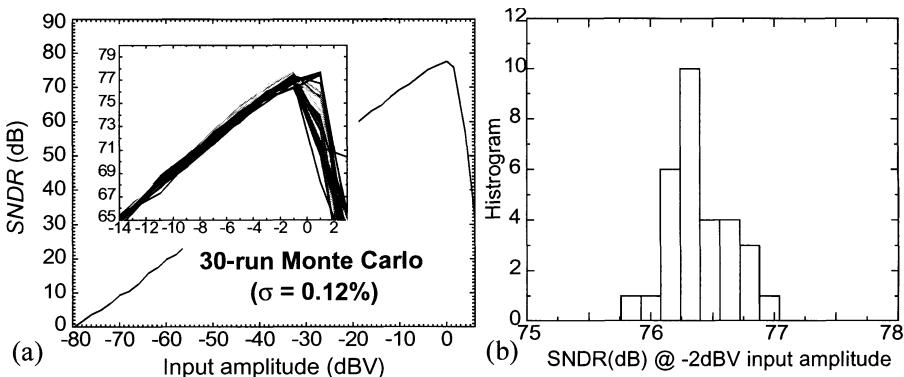


Figure 15.14: (a) Monte-Carlo analysis of the SNDR showing the capacitor mismatching effect. (b) Histogram of the SNDR for -2dBV input amplitude.

tion of the input amplitude simulated with ASIDES [1],[6]. This simulation is especially interesting for cascade structures due to their sensitivity to capacitor mismatch. Simulation takes 1s. for each iteration of the Monte Carlo analysis.

It is interesting to evaluate which is the accuracy of the behavioral simulation and how it compares with a conventional electrical simulator. ASIDES took 2.3s. to get the output power spectral density (computed from 65536 samples) shown in Fig. 15.15. Accuracy of this simulation can be verified by comparison with the experimental results (shown in the same figure) obtained from a chip prototype. A different signal frequency has been chosen for better visualization. Fig. 15.15 also shows the power spectral density (PSD) obtained with HSPICE. The electrical simulation took 5 days of CPU time to get only 8192 samples. It can be observed that HSPICE computes a lower error power because thermal noise cannot be included in the transient simulation.^{†1}

But there are other implementation possibilities, like using mathematical packages such as MATLAB. Also, the standardization of HDLs, such as VHDL [14] or Verilog and its analog extensions VHDL-AMS [15] and Verilog-A/MS, makes possible to use these languages and corresponding tools to model and simulate $\Sigma\Delta$ modulators. Moreover, these descriptions can be combined with HDL descriptions of other digital/analog/mixed-signal blocks. And this simulation is integrated in the design flow of commercial design environments. The main penalty is the simulation speed. The PSD plot in Fig. 15.15 was also computed by implementing the models in section 15.3.1 in VHDL and simulating with Mentor Graphics' Advance-MS. In this case, architectural model compilation takes 7.5s.

1. In all cases computation times were measured on a SunFire 3800 platform @ 750 MHz.

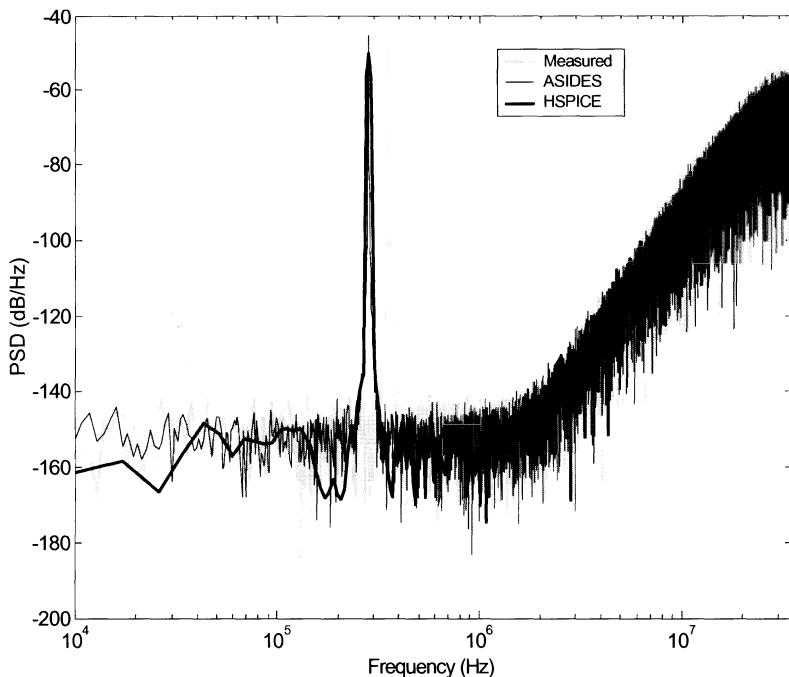


Figure 15.15: Simulated and measured PSD of the modulator in Fig. 15.2.

and simulation time takes 3.6s. The results are obviously identical to ASIDES because the same behavioral models were used.

Finally, we will compare the results of a worst-case simulation of the $\Sigma\Delta$ modulator in Fig. 15.2 by using ASIDES with the analytical models in section 15.3.1 and by building first a look-up table for the first two integrators of the modulator and then replacing the evaluation of the analytical models of these integrators by interpolated data from those tables. Fig. 15.16 compares the simulation results. Different signal frequencies were used for better visualization. It is interesting to notice that the behavioral simulation using some look-up tables correctly predicts a third-order harmonic that does not appear in the ASIDES level 1 simulation. This is due to the approximate analytical modeling of the opamp non-linearities. The use of the look-up tables makes computation time to increase from 2.3s. to 7.1s. Around 1 hour of CPU time to build the look-up tables using electrical simulation must be added to this time.

15.4 TOP-DOWN SYNTHESIS

The most knowledge-intensive processes of the hierarchical top-down method-

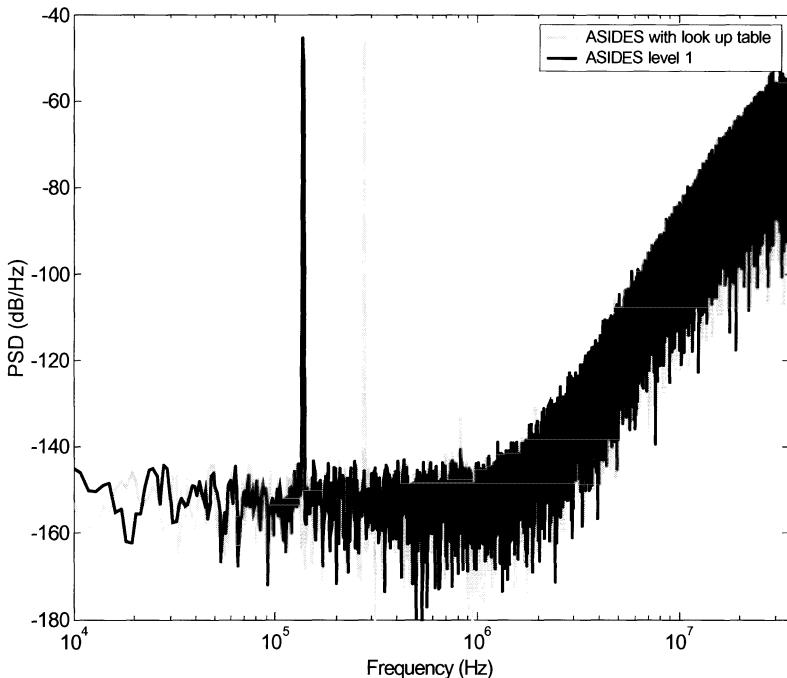


Figure 15.16: Illustrating look-up table vs. analytical behavioral simulation.

ology are the synthesis tasks. For the $\Sigma\Delta$ converter design, two synthesis levels can be identified in section 15.2: the modulator level and the cell level.

As a synthesis example all through this section, the design of a 14-bit 4MS/s $\Sigma\Delta$ modulator for xDSL applications in a $0.35\mu\text{m}$ CMOS technology will be considered.

15.4.1 Synthesis at the modulator level

15.4.1.1 Brief review of approaches

A few methodologies have been reported for high-level modulator sizing. Most of them, e.g., [16]-[18], perform what is called a knowledge-based synthesis, that is, expert knowledge is captured in the form of heuristic decisions or approximate equations. Sizing is carried out by following a design plan. They are in general topology- and technology-specific, e.g., [16] is limited to first- and second-order $\Sigma\Delta$ modulators, or [18] is limited to second and third-order modulators. In some cases, like [17], the lack of flexibility of knowledge-based tools is palliated by a deep hierarchical decomposition which enables the reuse of blocks through the

hierarchy and/or examine alternative topologies for sub-blocks.

More modern approaches use basically an iterative optimization procedure with a performance evaluator in the loop, as the generic diagram in Fig. 15.17 illustrates. At each iteration, circuit performances are evaluated at a given point of the design parameter space. According to such an evaluation, a movement in the design parameter space is generated and the process is repeated again. As it becomes clear from section 15.3, electrical simulation is prohibitive at the modulator level and, therefore, analytical equations or behavioral simulation are the only alternatives for performance evaluation at this level.

The methodology described hereinbelow imitates the common designers' procedures [1],[6]. Simple explicit equations are used for topology selection and rough sizing. Models get more complicated for certain candidate topologies and regions of the design space. And finally, detailed behavioral simulation together with new optimization procedures are executed for design fine-tuning. Another approach [7] omits the top-down refinement process and directly applies a behavioral simulator coupled with an optimization algorithm, specifically a type of genetic algorithm.

15.4.1.2 Optimization algorithm

As shown in Fig. 15.17, optimization procedures are an essential part of many synthesis systems and, in particular, of the $\Sigma\Delta$ modulator synthesis tool described herein. It becomes clear that the quality of the results (in terms of obtained performances, execution speed, cost, etc.) will critically depend on the features of the

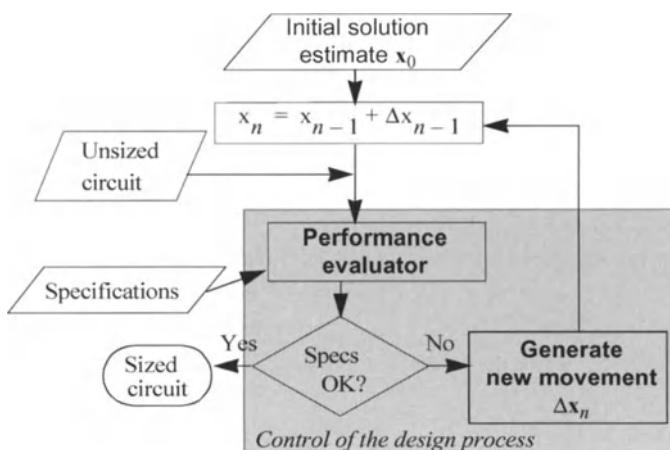


Figure 15.17: Optimization-based sizing.

optimization algorithms used, and therefore, they deserve some consideration.

The high-level (as well as the cell-level) synthesis of the $\Sigma\Delta$ modulator, like, in general, any electronic system, can be formulated as a constrained optimization problem. Unfortunately, the synthesis problem does not usually have an analytical solution. Therefore, a convenient approach is to use some kind of optimization procedure, which tries to minimize a cost function. This cost function use some kind of performance evaluator (based on equations or simulation) to quantify the degree of compliance of the performance specifications. There are basically two alternatives for the implementation of this iterative process:

- **Deterministic incremental techniques**, where parameter updating requires information on the cost function and on their derivatives. An important disadvantage is that only changes of design parameters that generate a decrease of the cost function are allowed. Therefore, the optimization process may be quickly trapped in a local minimum of the cost function. So, the usefulness of these techniques concentrates on the fine tuning of suboptimal sizings, that is, cases in which the initial solution in Fig. 15.17 is quite closed to the optimum one.
- **Statistical techniques**, where design parameters are varied randomly and, hence, information on the derivatives of the cost function is not required. The main advantage of the statistical techniques with respect to the deterministic ones is the capability to escape from local minima thanks to a non-zero probability of accepting movements that increase the cost function [19]. Therefore, these techniques are appropriate for global optimization, that is, cases in which no good initial solution in Fig. 15.17 is available. The price to pay is a larger computational cost.

A combination of both optimization techniques in two steps provides the best results: in the first one, statistical optimization techniques are applied, while deterministic ones are applied in the second step.

Some innovative features in the generation and acceptance of movements through the design parameter space allow to drastically reduce the computational cost of the statistical optimization techniques with respect to conventional simulated annealing algorithms; namely, exploration of the design space using a coarse grid to determine the best regions for finer exploration, adaptive control of the statistical optimization algorithm used (a kind of adaptive simulated annealing to synchronize design parameter update and the control parameter – named temperature – as a function of the percentage of accepted movements), etc. [1],[6].

15.4.1.3 *Synthesis system*

The synthesis system in [1],[6] is composed of architecture selection and com-

ponent sizing. Fast exploration of the design space and initial modulator sizing are based on equations and optimization techniques. These equations are basically of two types [1]:

- a) Architecture-specific equations representing the quantization noise power as a function of the non-idealities that affect its shaping function: integrator leakage, mismatching, etc.
- b) Equations related to other error sources: thermal noise, defective settling, jitter, integrator non-linearity, etc. It is assumed that the in-band error power is dominated by the first integrator because the contribution of the other integrators is attenuated by larger powers of the oversampling ratio.

Additionally, equations representing fundamental limits of the different topologies are also included.

Although development of this equation database implies an important effort, there are several reasons supporting this solution. On the one hand, many of these equations can be shared by many architectures. On the other, the design space is so vast and there are so many architectural choices that a more open solution based on general behavioral simulation would either take too much time or it would produce suboptimal solutions due to a deficient search of the design space.

Architecture selection requires an essential component in analog designers: knowledge. Therefore, the methodology presented herein includes tools intended to ease the acquisition of such knowledge. Some general considerations can be applied first for selecting a set of candidate architectures. For instance, a power-efficient $\Sigma\Delta$ modulator for the signal frequency required above, needs a low oversampling ratio. To attain a high resolution with such a moderate oversampling ratio, multi-bit quantization and high-order filtering are the possible strategies. In this context, cascade multi-bit architectures arise as an attractive choice. Quantization error and D/A nonlinearity are the dominant error sources in these architectures. In particular, dual quantization architectures, in which only the last stage has a multi-bit quantizer, are especially interesting. In this way, the errors associated to the multi-bit D/A converter can be largely attenuated at the modulator output.

Still, the choice of possible architectures is very large: oversampling ratio, order, number of bits in quantizers. The quantization noise should be attenuated in the low-frequency region with no distortion of the signal. Therefore, a set of coefficient relationships and digital functions can be easily obtained for each architecture, e.g., the relationships and functions in Table 15.1 result for the 2-1-1 multi-bit architecture in Fig. 15.2.

In principle, any set of coefficients satisfying the relationships in Table 15.1

should yield a working modulator. Nevertheless, for a real implementation the following considerations must be taken into account: (a) the signal level transferred from one stage to the following one must be low enough to avoid overloading of the latter; (b) the output swing required for the integrators must be physically achievable; and (c) the digital coefficient which amplifies the quantization error of the last stage must be as small as possible. The selection of the integrator weights becomes, in this way, an optimization problem which can be easily solved by a combination of a behavioral simulator (ASIDES in this case) and optimization algorithms.

Let us consider first the third order 2-1 cascade multi-bit structure. The in-band noise power at the modulator output due to quantization noise and the DAC-induced error is:

$$P_{2-1mb} \cong d^2 \left(\sigma_Q^2 \frac{\pi^6}{7M^7} + \sigma_D^2 \frac{\pi^4}{5M^5} \right) \quad (15.25)$$

where σ_Q^2 and σ_D^2 stand for the quantization error power and the D/A conversion error power, respectively, and M is the oversampling ratio. For the fourth-order 2-2 cascade multi-bit structure the in-band noise power is:

$$P_{2-2mb} \cong d^2 \left(\sigma_Q^2 \frac{\pi^8}{9M^9} + \sigma_D^2 \frac{\pi^4}{5M^5} \right) \quad (15.26)$$

and for the fourth-order 2-1-1 multi-bit architecture is:

$$P_{2-1-1mb} \cong d^2 \left(\sigma_Q^2 \frac{\pi^8}{9M^9} + \sigma_D^2 \frac{\pi^6}{7M^7} \right) \quad (15.27)$$

The DAC-induced error is attenuated by M^7 in the 2-1-1mb architecture

Table 15.1: Coefficient relationships and digital functions in Fig. 15.2.

Digital	Digital/Analog	Analog
$H_1(z) = z^{-1}$	$d_0 = 1 - g_3'(g_1g_2g_3)$	$g_1' = g_1$
$H_2(z) = (1 - z^{-1})^2$	$d_1 = g_3''/(g_1g_2g_3)$	$g_2' = 2g_1'g_2$
$H_3(z) = z^{-1}$	$d_2 = \left(1 - \frac{g_3'}{g_1g_2g_3}\right)\left(1 - \frac{g_4'}{g_3''g_4}\right) \equiv 0$	$g_4' = g_3''g_4$
$H_4(z) = (1 - z^{-1})^3$	$d_3 = g_4''/(g_1g_2g_3g_4)$	

whereas only by M^5 in the 2-1mb and 2-2mb architectures, which considerably increases the sensitivity of the latter ones to the DAC non-linearity. Therefore, these will not be further considered.

The optimization of integrator weights for the 2-1-1mb architecture yields an amplification of the quantization error of the last stage by $d_3 = 2$. This means a systematic loss of resolution of only 1 bit with respect to the ideal case. Another candidate architecture is the 5-th order 2-1-1-1mb topology. In this case, the attenuation of the DAC errors of the last stage grows up to M^9 whereas the systematic loss is kept at 1 bit. Although the attenuation of the DAC errors of the last stage is equal or larger in other architectures like the 2-2-1mb and the 2-2-2mb topologies, the set of coefficients obtained from the optimization procedure yield an amplification of the quantization error of the last stage by a factor 8, which means 3 bits of systematic loss. In conclusion, the 2-1-1mb and 2-1-1-1mb architectures remain as optimum ones after this first screening.

To select the optimum topology, behavioral simulation of these two topologies considering all other non-idealities is performed (see section 15.3). Among these, integrator leakage due to the finite open-loop opamp dc gain and integrator weight mismatching are the next most important non-idealities in cascade modulators. They contribute to extra in-band error power and may limit the useful resolution of the last quantizer. Fig. 15.18 shows the results of the behavioral simulation of both architectures yielding the effective resolution versus the number of bits of the last quantizer and for varying oversampling ratio. From the dynamic requirements of the integrators at each behavioral simulation it is possible to estimate the power consumption. From the modulators which meet the required resolution, the minimum power consumption is obtained with $M=16$ and $B=4$ for the 2-1-1mb cascade architecture and $M=14$ and $B=3$ for the 2-1-1-1mb architecture. As weight mismatching is a critical factor here, we use now the behavioral simulation to estimate the worst-case resolution of both optimum architectures as a function of the weight mismatch. The results are shown in Fig. 15.19. It can be seen that the sensitivity to weight mismatch is significantly higher in the 2-1-1-1mb architecture. In this case, the technology process considered is single-poly. Multi-metal capacitors exhibit a matching slightly larger than 0.1%, which makes the 2-1-1-1mb architecture less appropriate for this application in this technology process.

An optimization procedure is then started with the optimization kernel and the equation database for the selected architecture. The optimization results summarized in Table 15.2 are obtained in a few seconds of CPU time. This table also shows the different contributions to the in-band error power.

The requirements in Table 15.2 strictly apply only to the first integrator

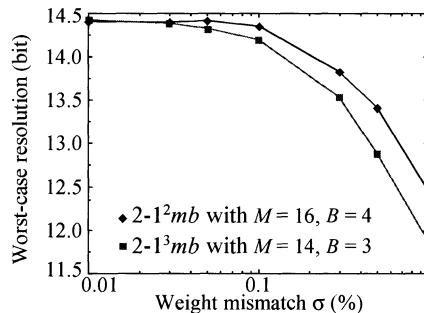


Figure 15.19: Simulated resolution as a function of the integrator weight mismatch ($A_v = 2500$, DAC INL = 0.4% FS).

because only that has been considered in the contributions to the in-band error power. The behavioral simulator ASIDES is then used in conjunction with the optimizer for fine-tuning of the specifications. In this way, specifications for the other opamps can be relaxed to avoid over-sizing and optimize the power consumption. The beneficial effect of this fine-tuning for the example above is patent in Table 15.3.

15.4.2 Synthesis at the cell level

15.4.2.1 Brief review of approaches

Reported approaches can also be classified at this level in two groups: knowledge-based approaches and optimization-based ones. As for high-level synthesis, knowledge-based approaches [16],[20] provide rough sizings through the use of a design plan. The execution time is very short but the quality of the results is usu-

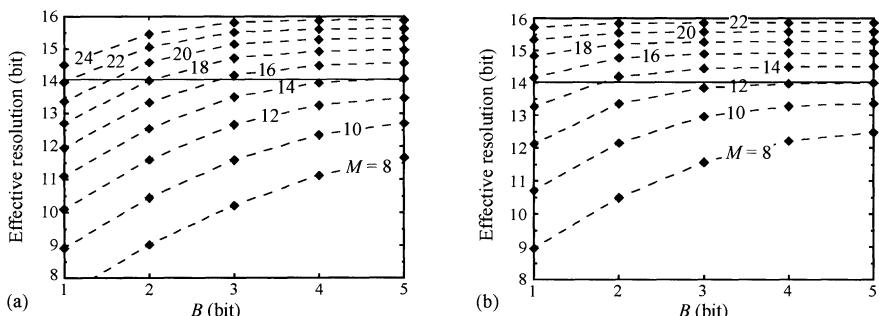


Figure 15.18: Modulator resolution vs. last-quantizer resolution for: (a) 2-1-1mb $\Sigma\Delta M$ and (b) 2-1-1-1mb $\Sigma\Delta M$ ($A_v = 2500$, weight mismatch $\sigma=0.1\%$, and DAC INL=0.4% FS).

Table 15.2: Modulator sizing results.

SPECS: 14bit@4MS/s@1.1V _p		2-1 ² mb
Modulator	Oversampling ratio	16
	Sampling frequency	64MHz
	Reference voltages	±2 V
Integrators	Sampling capacitor	0.5pF
	Unitary capacitor	0.5pF
	Sigma	0.12%
	Capacitor non-linearity ≤	25ppm/V
	Bottom parasitic capacitor	20%
	Switch ON-resistance ≤	250Ω
Opamps	DC-gain	68dB
	DC-gain non-linearity ≤	20%V ⁻²
	Transconductance ≥	2.5mA/V
	Maximum output current ≥	0.95mA
	Output swing ≥	±2 V
Comparators	Hysteresis ≤	20mV
A/D/A Converter	Resolution	4bit
	Non-linearity (<i>INL</i>) ≤	0.25%FS
Dynamic range		86dB 14bit
Quantization noise		-85.0dB
Thermal noise		-87.4dB
Incomplete settling noise		-99.7dB
Others		-106dB

Table 15.3: Opamp design parameters after fine-tuning.

	OA1		OA2	OA3	OA4
	Initial solution	ASIDES + FRIDGE	ASIDES + FRIDGE	ASIDES + FRIDGE	ASIDES + FRIDGE
gm	2.50mA/V	2.38mA/V	2.65mA/V	2.12mA/V	4.61mA/V
I _{max}	0.95mA	0.53mA	0.53mA	0.42mA	0.74mA
DCgain	2500	2500	1250	500	500

ally not good enough. These results are improved, in many cases, by the use of simulation tools, possibly combined with some local optimization. The lack of

flexibility, the technology-dependence, the high development and maintenance cost have made the interest in this kind of approaches to decrease.

Like in the high-level synthesis, optimization-based methods transform the sizing problem into a numerical optimization problem. Existing approaches differ in the numerical optimization algorithm and the method used to evaluate the cost function (the function which must be minimized in the optimization problem). According to this method we can distinguish equation-based and simulation-based approaches. In the former case, the value of the cost function is obtained from a set of analytical equations [21],[22]. The main advantage is the large speed for the cost function evaluation and, consequently, of the whole optimization process. The drawbacks are the approximate nature of the equations and the large effort to develop an appropriate set of equations for each topology (this effort is partially reduced in [22] by using a symbolic analysis tool).

Simulation-based approaches use electrical simulation for the evaluation of the cost function. This idea was first implemented in DELIGHT-SPICE [23]. However, the deterministic nature of the optimization algorithms limited the use to fine optimization of existing designs. The incorporation of statistical optimization enabled the sizing of basic cells starting from scratch [24],[1]. Subsequently, this same idea has been incorporated in other tools [25].

The FRIDGE tool [24],[1] uses the two-step global optimization algorithm, already described in section 15.4.1.2. Performance evaluation is performed through a modular interface to common electrical simulators, like HSPICE or Berkeley SPICE. This is enough to size any circuit described at a level understandable by the electrical simulators: any circuit that can be simulated can be sized. Nevertheless, it is common experience that during the synthesis process, designers generate pieces of knowledge that they wish to reuse in future designs. For this reason, FRIDGE also enables the incorporation of circuit-specific knowledge in the form of C procedures, which are compiled at run-time and used during the optimization procedure.

15.4.2.2 Synthesis example

When transmitting down the specifications of the building blocks, the results of the high-level synthesis of the $\Sigma\Delta$ modulator must be transformed into appropriate specifications of the basic cells. If we look, for instance, at the results for the integrators in Table 15.3, such parameters must be converted into appropriate design specifications of the amplifiers, e.g., gain-bandwidth product, slew-rate, etc., which depend on the loading conditions of each amplifier, which even change between the sampling and the integration phases. Equivalent loads for the sam-

pling and the integration phases of each amplifier must be calculated, yielding those in Table 15.4. From these and Table 15.3 the opamp specifications in Table 15.4 result.

Taking into account the specifications in Table 15.4, a two-stage OTA is selected for the first amplifier (Fig. 15.20(a)), whereas a folded-cascode OTA is used for the other three (Fig. 15.20(b)). FRIDGE automatically sizes each amplifier, meeting the specifications under worst-case simulation conditions. The performance specifications obtained from worst-case electrical simulation of the sized circuits are shown in Table 15.5.

Table 15.5: Simulated performances of the sized opamps.

	OA1	OA2	OA3	OA4	Unit
DC-gain	80.0	62.8	55.7	62	dB
GB	250 (1.6pF)	311 (1.4pF)	261 (1.4pF)	167 (4.5pF)	MHz
PM	67 (1.6pF)	65 (1.4pF)	71 (1.4pF)	79 (4.5pF)	°
Output swing	±2.5	±3.1	±2.9	±3.1	V
Slew rate	450	895	706	632	V/μs
Power consumption	38.5	4.5	4.0	6.6	mW

15.5 BOTTOM-UP LAYOUT CONSTRUCTION AND VERIFICATION

Once the converter has been sized at the modulator and cell level, it must be laid out. Basically, there exist two alternative approaches to layout synthesis.

One of them is called macrocell layout style. These methods use a set of procedural module generators for a set of basic layout primitives – macros–, typically single transistors or special groups of them, like matched transistors. Different geometrical variants usually exist for a given primitive, for each of which a module generator is needed. As such module generators are usually technology-spe-

Table 15.4: Equivalent loads and opamp specifications.

Amplifier	$C_{eq,s}$	$C_{eq,i}$	DC-gain	open-loop GB (load)	integrator SR
OA1	1.49pF	1.51pF	68dB	237MHz (1.6pF)	430V/μs
OA2	1.24pF	1.32pF	62dB	301MHz (1.4pF)	834V/μs
OA3	1.24pF	1.32pF	54dB	241MHz (1.4pF)	667V/μs
OA4	0.56pF	4.49pF	54dB	163MHz (4.5pF)	558V/μs

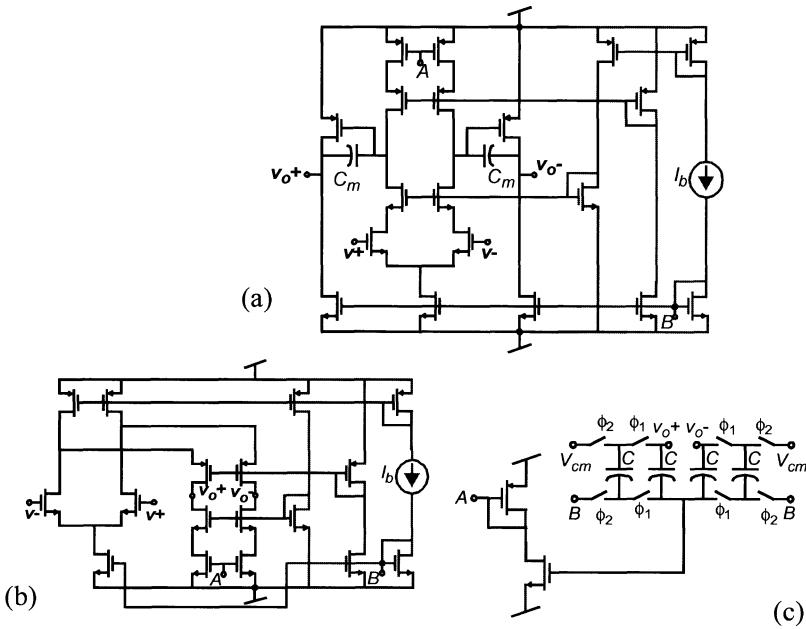


Figure 15.20: (a) Two-stage opamp; (b) folded-cascode opamp; (c) common-mode feedback circuitry for both amplifiers.

cific, maintenance is a costly process.

For given device sizes, the appropriate primitive variant must be selected and placed. Then, the cells are routed and, sometimes, a final compaction takes place. Placement and routing is formulated as a statistical optimization problem which must take into account typical analog constraints like minimization of parasitics, capacitive couplings, area, etc. In principle, they are very flexible methodologies, opened to any circuit architecture. The price to pay is a relatively high computation time. Besides, the formulation of a cost function for the optimization process which take into account all analog constraints is not a trivial task [26]-[28].

The second alternative is the use of layout templates, which store the relative position and interconnection of devices [29]. A major benefit of layout templates is that their instantiation for given device sizes is an extremely fast process. A major drawback of these methods is the high cost of template development, which can only be compensated through an intensive reuse of such layout templates. An obstacle for most approaches is that layout templates have to be rebuilt for each new technology process. Another drawback of these methods is the lack of flexibility, due to the difficulties to efficiently accommodate widely varying device

sizes.

The methodology in [30] palliates above drawbacks. On the one hand, it introduces a technology-independent methodology for construction of layout templates based on commercial tools. On the other hand, layout quality in terms of performance and area occupation is evaluated at each iteration of the optimization process. This layout quality plays a similar role to the evaluation of other circuit performances. The result of the optimization process becomes, in this case, a sized circuit which tries not only to meet the design specifications but also to yield a good quality layout.

The use of layout templates has an additional advantage: parasitics can be easily estimated during the iterative synthesis procedures without template instantiation. This coupling of synthesis and layout phases minimizes the number of re-design iterations in the practical application of the methodology in Fig. 15.1.

For illustration's sake, Table 15.6 collects the results of three sizing experiments using the FRIDGE tool. The circuit to size is the fully-differential two-stage operational amplifier in Fig. 15.21. In the first experiment, neither layout quality evaluation nor parasitic estimates are considered during the size tuning process. Area minimization is pursued by assuming a direct relationship with device sizes and not using any information from the layout template. In the second experiment, layout quality is evaluated but the only parasitic estimates are those obtained directly from device sizes. In the third experiment, both, parasitic estimates and evaluation of layout quality, are taken into account. In the three experiments $V_{DD} = 1.5V$, $V_{SS} = -1.5 V$, $C_L = 8 \text{ pF}$ and $R_L = 100 \text{ kW}$. The second column in Table 15.6 displays the required specification values for the performance parameters listed in the first column. Third, fourth and fifth columns show the results of each experiment and the numbers between brackets correspond to the specification values obtained after layout template instantiation, extraction and simulation.

Template instantiation of the resulting devices sizes are shown in Fig. 15.22 (notice that all layouts have been captured with the same resolution and, therefore, the relative dimensions are real). It can be seen in Table 15.6 that the required nominal performance is accomplished in the three experiments. But, the exclusion of layout quality evaluation and/or parasitic estimates in the sizing process, results in additional re-design iterations (when, due to layout parasitics, the final design does not meet the required specifications - see phase margin of experiments I and II in Table 15.6) or in hardly compact layouts with large empty areas (when the sizing process is geometrically unconstrained).

Coming back to the synthesis example of the previous section, Fig. 15.23

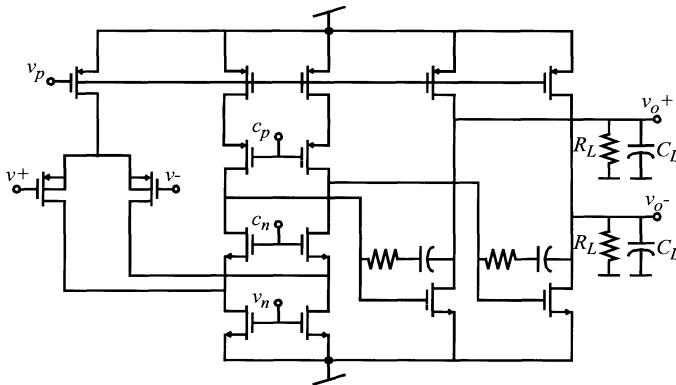


Figure 15.21: Fully-differential Miller-compensated two-stage amplifier.

shows the layout of the $\Sigma\Delta$ modulator. Final verification prior to fabrication requires simulation of the extracted layout taking into account process variations, supply tolerances and temperature ranges. In the present case, the number of necessary simulations grows to 110 corners. From the discussions in section 15.3, it becomes clear that electrical simulation of the extracted layout is only feasible for a single operating condition and a reduced number of samples. Highest accuracy would be provided by using look-up tables for the behavioral simulation of the complete modulator. However, it has been shown that building the look-up tables is a costly process and different tables are needed for each corner simulation. A more convenient approach is to update the parameters of the behavioral models in ASIDES with data extracted from electrical simulation. In a first step, an ASIDES

Table 15.6: Experimental results.

Spec	Value	Experiment I	Experiment II	Experiment III	Units
DC gain	>110	110.0 (110.0)	112.6 (112.6)	113.0 (113.0)	dB
GB	>90	91.8 (90.0)	98.6 (91.9)	105.7 (105.4)	MHz
Phase margin	>65	67.6 (63.4)	65.3 (61.9)	65.4 (65.3)	deg
Output swing	>5.25	5.3 (5.3)	5.4 (5.4)	5.3 (5.3)	V
Slew rate	>40	46.9 (46.8)	41.0 (40.4)	57.2 (57.0)	V/ μ s
Inp. noise@1kHz	<40	34.0 (34.0)	36.9 (37.0)	30.5 (30.5)	(nV)/($\sqrt{\text{Hz}}$)
Power dissipation	minimize	7.3 (7.3)	4.2 (4.2)	8.0 (8.0)	mW
Total area	minimize	195.8 \times 358.8	190.8 \times 173.6	173.8 \times 191.25	μm^2
Aspect ratio	≈ 1	0.55	1.1	0.9	-----

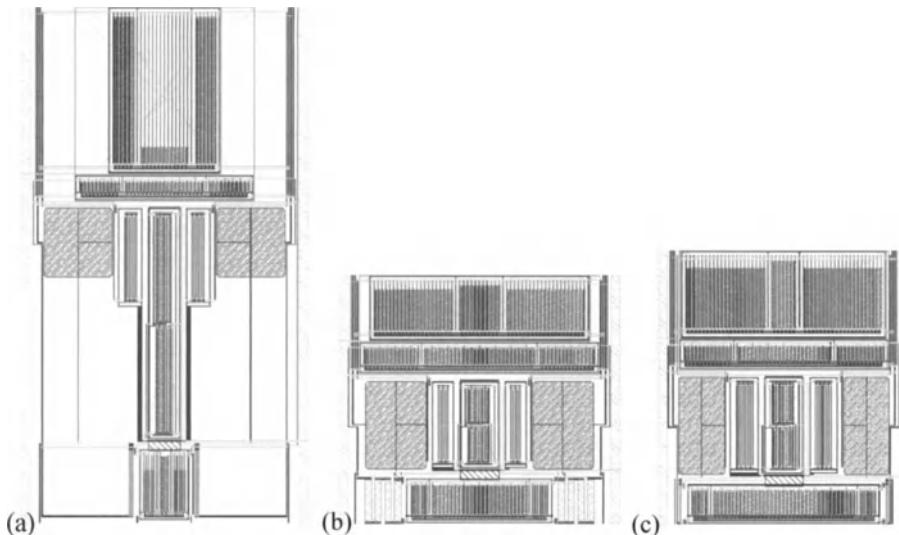


Figure 15.22: Layout instantiation for the three experiments in Table 15.6.

simulation is performed in every corner. In a second step, look-up tables are built for the worst-case corner results and simulation is repeated using these look-up tables. This worst-case simulation for the extracted layout from Fig. 15.23 provides an integrated band noise of -83dB from DC to 2 MHz, which corresponds to a dynamic range of 86 dB (equivalently 14 bits).

15.6 REFERENCES

- [1] F. Medeiro, B. Pérez-Verdú and A. Rodríguez-Vázquez, *Top-Down Design of High-Performance Sigma-Delta Modulators*. Kluwer, 1999.
- [2] G.E. Gielen and J.E. Franca, “CAD Tools for Data Converter Design: An Overview”, *IEEE Trans. on Circuits and Systems II*, Vol. 43, No. 2, pp. 77-89, Feb. 1996.
- [3] V. F. Dias, V. Liberali and F. Maloberti, “Design Tools for Oversampling Data Converters: Needs and Solutions”. *Microelectronics Journal*, Vol. 23, pp. 641-650, 1992.
- [4] C. H. Wolff and L. Carley: “Simulation of Δ - Σ Modulators Using Behavioral Models”, *Proc. IEEE Int. Symp. on Circuits and Systems*, pp. 376-379, 1990.
- [5] V. Liberali, V.F. Dias, M. Ciapponi and F. Maloberti, “TOSCA: a Simulator for Switched-Capacitor Noise-Shaping A/D Converters,” *IEEE Trans. Computer-Aided Design*, Vol. 12, No. 9, pp. 1376-1386, Sept. 1993.
- [6] F. Medeiro, B. Pérez-Verdú, A. Rodríguez-Vázquez and J.L. Huertas, “A Vertically Integrated Tool for Automated Design of SD Modulators,” *IEEE J. Solid-State Circuits*, Vol. 30, No. 7, pp. 762-772, July 1995.

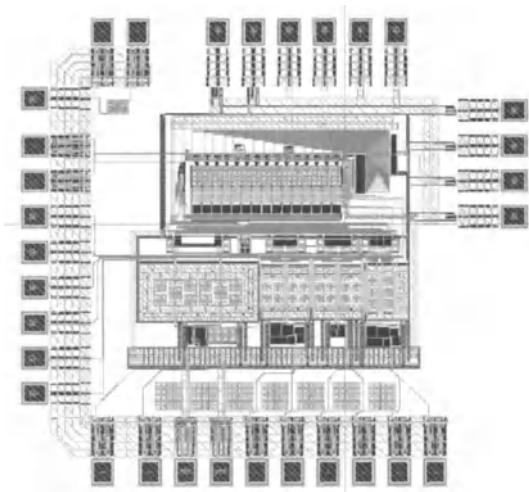


Figure 15.23: Layout of the cascade modulator.

- [7] K. Francken, P. Vancorenland and G. Gielen, “DAISY: a simulation-based high-level synthesis tool for $\Delta\Sigma$ modulators,” *Proc. IEEE/ACM Int. Conf. Computer-Aided Design*, pp. 188-192, 2000.
- [8] G.T. Brauns, R.J. Bishop, M.B. Steer, J.J. Paulos and S.H. Ardalan, “Table-based Modeling of delta-sigma modulators using ZSIM,” *IEEE Trans. Computer-Aided Design*, pp. 142-150, Vol. 9, No. 2, Feb. 1990.
- [9] R. del Río, F. Medeiro, B. Pérez-Verdú and A. Rodríguez-Vázquez, “Reliable analysis of settling errors in SC integrators-application to the design of high-speed $\Sigma\Delta$ modulators,” *Proc. IEEE Int. Symp. Circuits and Systems*, Vol. 4, pp. 417-420, 2000.
- [10] R. del Río, F. Medeiro, B. Pérez-Verdú and A. Rodríguez-Vázquez, “Reliable analysis of settling errors in SC integrators: application to $\Sigma\Delta$ modulators,” *IEE Electronic Letters*, Vol. 36, No. 6, pp. 503-504, Mar. 2000.
- [11] K. Lee and R. G. Meyer, “Low Distortion Switched-Capacitor Filter Design Techniques”, *IEEE Journal of Solid-State Circuits*, Vol. 20, pp. 1103-1113, Dec. 1985.
- [12] C.B. Wang, “A 20-bit 25-kHz Delta-Sigma A/D Converter Utilizing a Frequency-Shaped Chopper Stabilization Scheme,” *IEEE J. Solid-State Circuits*, Vol. 36, No. 3, pp. 566-569, March 2001.
- [13] J.W. Bennett: “Spectra of Quantized Signals”, *Bell Syst. Tech. J.*, Vol. 27, pp. 446-472, July 1948.
- [14] *IEEE VHDL Language Reference Manual*, IEEE Std 1076-2002.
- [15] *IEEE VHDL 1076.1 Language Reference Manual*, IEEE Std 1076.1-1999.
- [16] M. Degrauwé et al., “IDAC: An Interactive Design Tool for Analog CMOS Circuits,” *IEEE J. Solid-State Circuits*, Vol. 22, pp. 1106-1114, Dec. 1987.
- [17] G. Beenker, J.D. Conway, G. Schrooten and A. Slenter, “Analog CAD for Consumer

- IC's," *Proc. Workshop on Advances in Analog Circuit Design*, pp. 343-355, 1992.
- [18]M.F. Mar and R.W. Brodersen, "A Design System for On-Chip Oversampling A/D Interfaces," *IEEE Trans. on VLSI Systems*, Vol. 3, pp. 345-354, Sept. 1995.
- [19]P. J. M. Laarhoven and E.H.L. Aarts: "Simulated Annealing: Theory and Applications", Kluwer, 1987.
- [20]R. Harjani, R. Rutenbar and L.R. Carley: "OASYS: A Framework for Analog Circuits Synthesis", *IEEE Trans. Computer-Aided Design*, Vol. 8, pp. 1247-1265, December 1989.
- [21]H. Y. Koh, C.H. Sequin and P.R. Gray: "OPASYN: A Compiler for CMOS Operational Amplifier", *IEEE Trans. Computer-Aided Design*, Vol. 9, pp. 113-125, February 1990.
- [22]G. Gielen, H. Walscharts and W. Sansen, "Analog Circuits Design Optimization Based on Symbolic Simulation and Simulated Annealing", *IEEE J. Solid-State Circuits*, Vol. 25, pp. 707-713, June 1990.
- [23]W. Nye, D.C. Riley, A. Sangiovanni-Vincentelli and A.L. Tits, "DELIGHT.SPICE: An Optimization-Based System for the Design of Integrated Circuits", *IEEE Trans. Computer-Aided Design*, Vol. 7, pp. 501-519, April 1988.
- [24]F. Medeiro, R. Rodríguez-Macías, F.V. Fernández, R. Domínguez-Castro, J.L. Huertas and A. Rodríguez-Vázquez, "Global Design of Analog Cells Using Statistical Optimization Techniques", *Analog Integrated Circuits and Signal Processing*, Vol. 6, No. 3, pp. 179-195, Kluwer, Nov. 1994.
- [25]R. Phelps, M. Krasnicki, R.A. Rutenbar, L.R. Carley and J.R. Hellums, "Anaconda: Simulation-Based Synthesis of Analog Circuits Via Stochastic Pattern Search," *IEEE Trans. Computer-Aided Design*, Vol. 19, No. 6, pp. 703-717, June 2000.
- [26]J. Rijmenants, et al., "ILAC: An Automated Layout Tool for Analog CMOS Circuits," *IEEE J. Solid-State Circuits*, Vol. 24, pp. 417-425, April 1989.
- [27]J.M. Cohn, R.A. Rutenbar and L.R. Carley, "KOAN/ANAGRAM II: New Tools for Device-Level Analog Placement and Routing." *IEEE J. Solid-State Circuits*, Vol. 26, No. 3, pp. 330-342, March 1991.
- [28]K. Lampaert, G. Gielen and W. Sansen, "A Performance-Driven Placement Tool for Analog Integrated Circuits." *IEEE J. Solid-State Circuits*, Vol. 30, No. 7, pp. 773-780, July 1995.
- [29]J.D. Conway and G.G. Schrooten, "An Automatic Layout Generator for Analog Circuits," *Proc. IEEE European Design Automation Conf.*, pp. 513-519, 1992.
- [30]R. Castro-López, F.V. Fernández, F. Medeiro and A. Rodríguez-Vázquez, "Generation of Technology-Independent Retargetable Analog Blocks," *Analog Integrated Circuits and Signal Processing*, Vol. 33, No. 2, pp. 157-170, Kluwer, Nov. 2002.

Chapter 16

SYSTEMATIC COMPUTER-AIDED DESIGN METHODOLOGY FOR CMOS DATA CONVERTERS

Analog design in the footsteps of Antonio Gaudi

Georges Gielen and Kenneth Francken

Abstract

This chapter presents a systematic design methodology for CMOS data converters that are used as embedded macrocells in integrated VLSI systems. The approach is illustrated with examples from different types of data converters. A generic behavioral model, that includes the most important nonidealities in a parameterized way, is used for system-level exploration to define the converter's individual specifications within the system. The architecture of the converter and the sizes of the devices are then determined using a two-level performance-driven design methodology. At the converter architectural level the architectural decisions (the exact converter topology and the subblock requirements) are determined. This can be automated by combining an optimization algorithm with a fast behavioral simulation of the converter. The derived subblock specifications then serve as input for the circuit sizing (optimization) of the individual subblocks. This is then followed by the layout generation phase. This procedure will be illustrated with the high-level synthesis of a $\Delta\Sigma$ converter and the entire design of a Nyquist-rate converter. Also, special layout tools for the generation of regular structures with complex connectivity, typical for data converters, are presented. Finally in the design methodology, a detailed behavioral model is extracted from the completed design, combining in accurate detail the major nonidealities such as static (e.g. INL, DNL) and dynamic (e.g. glitch, SFDR) behavior. The whole methodology is illustrated with several data converter design examples. The experimental results demonstrate how the described methodology and supporting tools drastically reduce the total design time for embedded data converters, thereby significantly increasing analog design productivity.

16.1 INTRODUCTION

In these days the design of heterogeneous microelectronic systems on a single or a few chips is becoming feasible due to the ever decreasing feature size of the CMOS silicon technology. These systems implement functions that require both digital blocks (DSPs, microprocessors, reconfigurable digital blocks), memories (RAM/ROM/EEPROM) as well as analog macrocells (D/A and A/D converters, drivers, frontends, etc.). The use of cores and other IP (intellectual property) blocks as well as new methodologies such as platform-based design promise the design

productivity boost that is needed to generate these systems in the shortening time to market constraints. The design of analog functional blocks, however, requires a large amount of effort compared to digital or DSP systems, especially when large and complex blocks as for instance high-accuracy converters are considered.

To tackle this problem, a number of approaches have been proposed. Analog synthesis tools promise an automated solution for the design of analog building blocks [1]. These tools use a top-down, constraint-driven design methodology to refine system-level block specifications into a fully completed transistor-level design. The application range of most analog synthesis tools today, however, is limited to usually one or a few types of circuits [2–5], and only few approaches cover the complete design flow from specifications down to layout. Nevertheless, full automation of the design of analog macrocells is only useful for those types of blocks that show a high level of topological reuse. Examples are $\Delta\Sigma$ converters, pipelined converters, etc. For the more high-performance designs with ever changing topological variations, by definition no automation can be provided. Nonetheless a tool-supported systematic design methodology can also drastically reduce the design time for these cutting-edge designs as will also be illustrated in this chapter for several data converter examples. Like in the work of Antonio Gaudi, systematic design engineering and tool support do not necessarily come at the expense of performance and beauty.

In this chapter a systematic computer-aided design methodology for high-performance CMOS data converters is proposed and demonstrated for several converter design examples. The design methodology covers the complete design flow and is supported by software tools to speed up the task significantly. A generic behavioral model, that includes the most important nonidealities in a parameterized way, is used for system-level exploration to define the converter's individual specifications within the system. The architecture of the converter and the sizes of the devices are then determined in the selected technology process using a two-level performance-driven design methodology. At the converter architectural level the architectural decisions (the exact converter topology and the subblock requirements) are determined. This can be automated by combining an optimization algorithm with a fast behavioral simulation of the converter. The derived subblock specifications then serve as input for the circuit sizing (possibly again by optimization) of the individual subblocks. This is then followed by the layout generation phase. This procedure will be illustrated with the high-level synthesis of a $\Delta\Sigma$ converter and the entire design of a Nyquist-rate converter. Also, special layout tools for the generation of regular structures with complex connectivity, typical for data converters, are presented. Finally, in the design methodology, a detailed behavioral model is extracted from the completed design, combining in accurate detail the major non-

idealities such as static (e.g. INL, DNL) and dynamic (e.g. glitch, SFDR) behavior. The results of the design methodology are therefore a generated layout and the corresponding extracted behavioral model that allows to include the converter in system simulations for final verification.

The systematic design method allows to speed up the generation of new designs for given specifications, or to easily port designs to new technology processes. The whole methodology is illustrated for several different data converter design examples. The experimental results demonstrate how the described methodology and supporting tools drastically reduce the total design time for embedded data converters, thereby significantly increasing analog design productivity. The presented methodology, although not necessarily fully automated, offers a large flexibility of the topology used, in this way covering a broad range of performance requirements.

The chapter is organized as follows. Section 16.2 explains the proposed systematic computer-aided design methodology. Section 16.3 presents some examples of behavioral models used for system-level design. The actual design of the converter blocks, both at the architectural and at the circuit level, is then explained to full extent in section 16.4. Also some practical design examples are given here. The section also describes the layout generation process with the Mondriaan tool. Section 16.5 presents the generation model. Finally, section 16.6 provides conclusions.

16.2 SYSTEMATIC DESIGN FLOW FOR DATA CONVERTERS

In the design of analog functional blocks as part of a large system on silicon (SoC), a number of different phases are identified as depicted in figure 16.1. The first phase in the design is the specification phase. During this phase the analog functional block is analyzed in relation to its environment, the surrounding system, to determine the system-level architecture and the block's required specifications. With the advent of analog hardware description languages such as VHDL-AMS or VERILOG-A/MS, the obvious implementation for this phase is a generic analog behavioral model. This model has to be parameterized with respect to the specifications of the functional block. Generic here means that no details about the exact implementation of the converter are known at this stage yet. The next phase in the design procedure is the actual design (synthesis) of the functional block (see center of figure 16.1), consisting of sizing and layout. The design methodology used here is top-down performance-driven [6]. This design methodology has been accepted as the de facto standard for systematically designing analog building blocks [1]. Since data converters are mixed analog-digital circuits, both the analog and the digital parts need to be designed. The analog design flow is grouped on the left in the center of figure 16.1, the corresponding digital flow is grouped on the right. The analog flow consists of a sizing at two hierarchical levels: the architectural level and the circuit

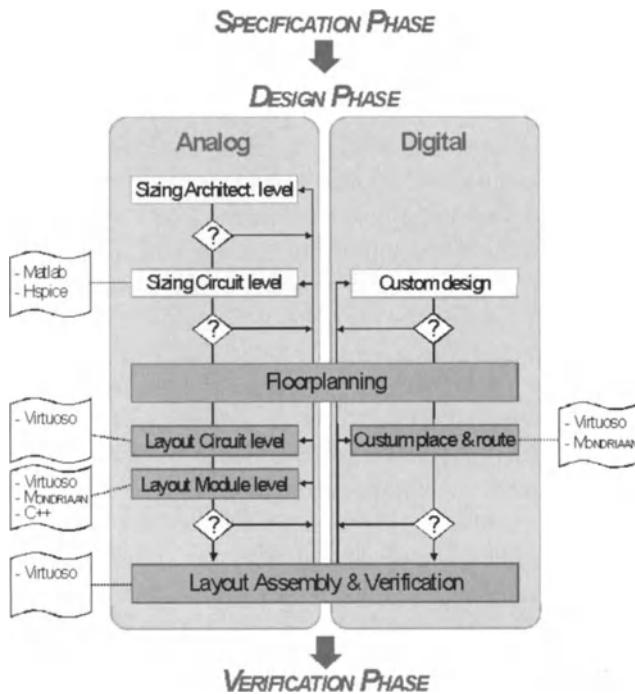


Figure 16.1: Presented systematic computer-aided design flow for high-performance data converters.

level. The digital synthesis completes the sizing part of the mixed-signal design. The design steps are verified using classical approaches (numerical verification with a simulator, at the behavioral, device and gate level respectively). The floorplanning is done jointly for the entire data converter including both the analog and digital blocks, after which the analog layout is generated, and standard cell place and route is used to create the digital layout. Both layouts are separately verified. The blocks are assembled at the module level and again a module-level verification is done with classical tools. When the full converter design is finished and verified in the final verification phase, the complete system in which the functional block is embedded, must be verified in the final verification phase, as shown at the bottom in figure 16.1. For this again a behavioral model for the analog functional block is constructed, but this time a more detailed model based on the actual circuit implementation can be used. The actual model parameters extracted from the generated layout are then used to verify the functioning of the block within the system.

The performance of Nyquist-rate converters is restricted by the trade-off between speed, power and accuracy [7]. The accuracy is limited by the mismatch between transistors. To improve the accuracy, larger devices are required that have better matching, but at the same time the capacitive loading on the circuit nodes increases and more power is required to attain a certain speed performance. Hence the fundamental trade-off:

$$\frac{\text{speed} \cdot \text{accuracy}^2}{\text{power}} \approx \frac{1}{C_{ox} \cdot A_{VT}^2} \quad (16.1)$$

where C_{ox} is the unit-area oxide capacitance and A_{VT} the MOS transistor threshold voltage mismatch parameter. This relationship implies that for today's circuits, that aim at high speed, high accuracy and low power, a technological limit is encountered in the mismatch of the devices. Therefore, the handling of statistical mismatch errors as well as any systematic errors that limit accuracy is key in any design flow for high-performance Nyquist-rate converters as will be illustrated below. Oversampling converters overcome this problem by converting the oversampling margin into increased resolution, at the expense of lower signal frequencies.

The remainder of the chapter focuses on the generic behavioral modeling, the architecture-level and circuit-level sizing synthesis, the layout generation and the behavioral model extraction steps in the design flow for high-performance CMOS data converters, and illustrated this with practical design examples.

16.3 GENERIC BEHAVIORAL MODELING FOR THE TOP-DOWN PHASE

By using a complete mixed-signal hardware description language model of the embedded block, the designer can explore different solutions at the system level in terms of performance, power and area consumption. In this way the high-level specifications of the system can be translated into individual specifications for the embedded data converter, as well as for the other blocks in the system. This approach is illustrated here with the generic behavioral model of both an analog-to-digital and a digital-to-analog Nyquist-rate data converter.

For a Nyquist-rate data converter the static input-output relationship (from analog input to digital output code for an A/D converter and from digital input code to analog output for a D/A converter) can be fully described by the transition points in the transfer curve of the converter (see figure 16.2). These transition points can be nonideal and can have a statistical spread due to mismatches. Hence, a quite generic model to characterize the static input-output behavior of a converter is to represent the converter by the vector of transition points according to [8]:

$$t \sim \text{normal}(\mu_t, \Sigma_t) \quad (16.2)$$

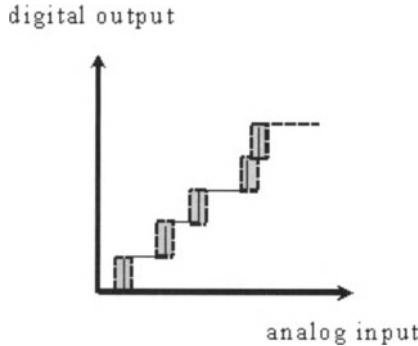


Figure 16.2: Static transfer characteristic of a Nyquist–rate data converter.

where μ_t contains the deterministic part and Σ_t the statistical part of the transfer characteristic. Effects like INL and DNL are included here, and random sample generation can be used to evaluate the effect of mismatches.

Besides static behavior also the dynamic effects have to be modeled. For the dynamic (transient) behavior of a D/A converter, the most important specification is the glitch energy. The response of a D/A converter to a sudden change in the input code typically looks as shown in figure 16.3, which includes both a finite settling and a ringing behavior. A generic model of the glitch can therefore be obtained by superposition of an exponentially damped sine and a shifted hyperbolic tangent [9]:

$$i_{out} = A_{gl} \cdot \sin\left(\frac{2\pi}{t_{gl}}(t - t_0)\right) \cdot \exp\left(-sgn(t - t_0)\frac{2\pi}{t_{gl}}(t - t_0)\right) + \frac{level_{i+1} - level_i}{2} \cdot \tanh\left(\frac{2\pi}{t_{gl}}(t - t_0)\right) + \frac{level_{i+1} + level_i}{2} \quad (16.3)$$

in which i_{out} is the output current, A_{gl} is the amplitude and t_{gl} the period of the glitch signal, and $level_i$ and $level_{i+1}$ are the code levels between which the converter switches. The glitch energy is defined as the integrated area indicated in gray in figure 16.3. Using this generic model and assigning proper values to the model parameters, the required specifications for the D/A converter can be derived by performing exploratory simulations at the system level.

16.4 SIZING SYNTHESIS OF THE DATA CONVERTER

The specifications that have been derived during the top–down specification phase, are now input to the sizing synthesis of the converter itself. The design of the converter is performed hierarchically [6], as indicated in figure 16.1, over

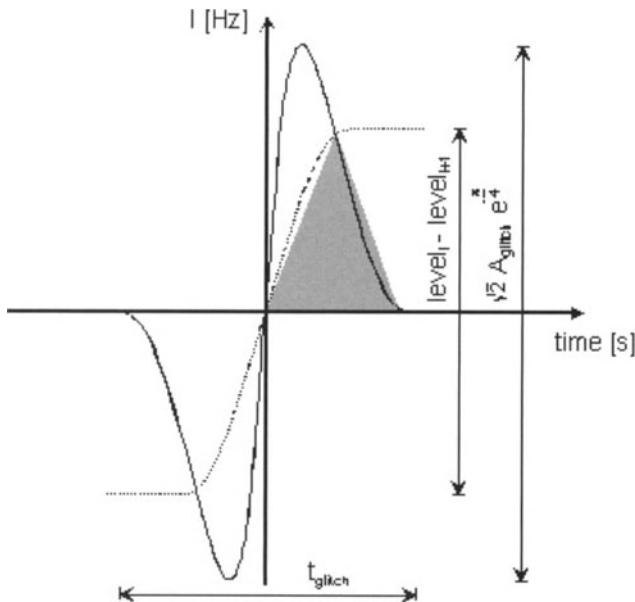


Figure 16.3: Transient response of a D/A converter when switching the input between two digital code levels.

two levels: the architecture level and the circuit level. First the high-level synthesis at the architectural level has to be performed, where both some architectural decisions are taken in order to finalize the converter architecture and secondly the minimum specifications for the building blocks within the converter are determined. These specifications then serve as input for the circuit-level sizing of these building blocks down to the transistor level.

Both phases could be performed manually or could be automated. When using a computer-aided formulation for either the architectural-level or the circuit-level sizing, most CAD tools today use an optimization-based formulation, where the sizing problem is cast as a constrained optimization problem in which the goal is to find the design solution that meets the specification requirements at minimal implementation cost (e.g. minimum power consumption) [1]. This is schematically illustrated in figure 16.4. The only differences between high-level architectural synthesis and circuit-level synthesis are that in the first case the optimization variables are performance requirements of the building blocks within the converter, that the simulations have to be performed at the behavioral level using behavioral mod-

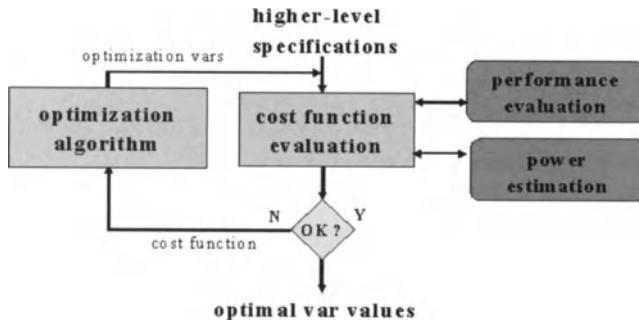


Figure 16.4: Schematic flow of optimization-based high-level architectural and circuit-level sizing of data converters and analog blocks in general.

els to represent the building blocks, and that the implementation cost is estimated using e.g. power estimators, whereas in the latter case the optimization parameters are sizes and biasing of the devices within the building block circuit schematics and the simulations are performed at the transistor level with a SPICE simulator. This two-phase methodology is now illustrated for two converter design examples.

16.4.1 Architectural-level synthesis of a $\Delta\Sigma$ A/D converter

The behavioral modeling of $\Delta\Sigma$ modulators is described quite extensively in literature (see [4, 10] for more details and references). The behavioral abstraction level renders the simulation of these oversampling converters feasible within a reasonable timeframe. In fact, the speed advantage is so high that a simulation-based optimization approach can be used for automatically sizing the converter at the architectural level. In the following subsection, we will discuss the methodology for efficiently tackling the high-level synthesis problem of $\Delta\Sigma$ modulators. By carefully modeling all major building blocks a sufficiently *accurate* result can be obtained as is shown in the second subsection, where a $\Delta\Sigma$ modulator satisfying the ADSL specifications is synthesized.

16.4.1.1 Methodology

The simulation-based synthesis uses the dedicated behavioral $\Delta\Sigma$ simulation tool DAISY [11] in combination with an optimization algorithm. The complete flow is visualised in figure 16.5. All interfaces between the modules are integrated in the DAISY tool as to ensure a completely automated synthesis run. More information

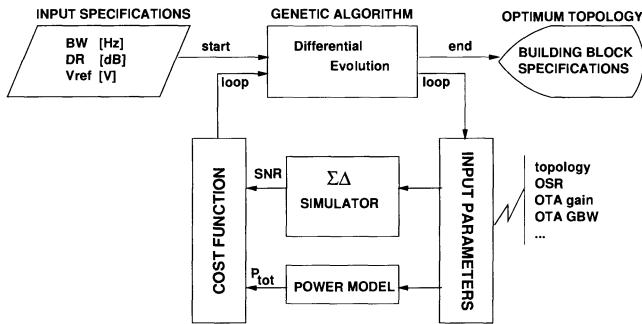


Figure 16.5: Flow of the high-level synthesis procedure for $\Delta\Sigma$ modulators.

on the simulation engine can be found in [10]. We will now discuss the remaining blocks of the synthesis loop.

Input / output

The typical performance specifications, namely dynamic range (in dB) and signal bandwidth (in Hz), serve as inputs to the tool. The user can also specify any other input (reference voltage, sampling frequency, ...) or nonideality (opamp gain, comparator offset, ...) that is modeled in the simulator. All these parameters can be fixed to a specific value, kept ideal or be included in the optimization process. Also technology-dependent data such as the minimum capacitor value can be set. After completing optimization, the tool then returns the optimum modulator topology with the lowest power consumption to achieve these specifications as well as the required specifications for all of the building blocks in the determined modulator topology.

Genetic algorithm

As optimization algorithm we employ the differential evolution algorithm used in [12], which we altered to fit our needs. It is a genetic algorithm that searches for a global optimum. Both the topology (architecture, oversampling ratio OSR, ...) and the building block specifications are optimized simultaneously.

Optimization parameters

The optimization parameters can be divided in parameters that control the topology, such as the type of the modulator structure and the oversampling ratio, and parameters that represent the building block specifications such as the OTA gain, GBW, output swing, the finite switch on-resistance, comparator offset and hys-

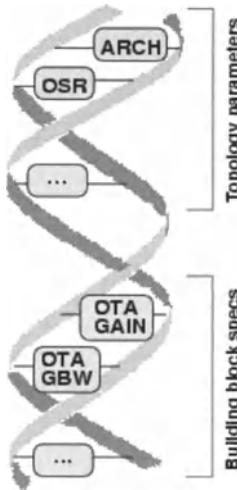


Figure 16.6: Representation of a population member in the genetic algorithm.

teresis, One population member ('chromosome') in the genetic algorithm is therefore represented as shown in figure 16.6. These parameters are passed to the simulator together with the signal bandwidth specification, so that the simulator can determine the correct sampling frequency (being two times the signal bandwidth times the OSR). Note that the oversampling ratio can itself be fixed or set as an optimization variable. The simulator then calculates the value of the dynamic range using time-domain simulations and FFT calculation. Also, the sensitivities of the modulator performance to the building block specifications are derived which are needed to make the design robust. The power model uses the same optimization variables to estimate the power consumption of the entire modulator as needed to find the minimum-power solution.

Cost function formulation

As in most optimization problems, the formulation of the cost function is crucial. The first part of our cost function is a penalty term proportional to the absolute value of the relative (linear) deviation of the simulated and specified dynamic range specification:

$$cost_1 = K_1 \cdot \text{abs} \left(\frac{10^{DR_{sim}} - 10^{DR_{spec}}}{10^{DR_{spec}}} \right) \quad (16.4)$$

where K_1 is a constant weight depending on whether the specification was met (1) or not (1E6). This is graphically illustrated in figure 16.7.

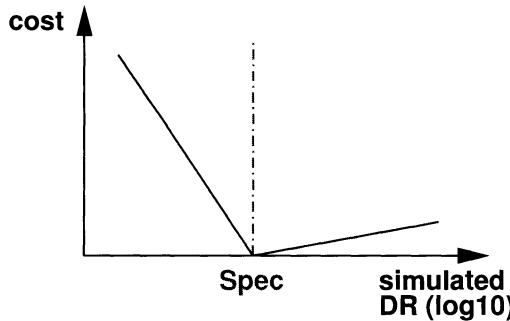


Figure 16.7: Cost penalty for deviations from the specified dynamic range (DR).

A second part of the cost function takes the relative power consumption into account:

$$cost_2 = K_2 \cdot P_{rel} \quad (16.5)$$

where K_2 is another constant weight that is set to 10^{11} based on experimental results. Relative means that we are not interested in the absolute value of the power but only in how the power changes when the optimization variables vary. The power model will be described later on.

Finally, the total cost is given by:

$$cost = cost_1 + cost_2 \quad (16.6)$$

It is, however, also possible that the genetic algorithm proposes bad combinations of parameters (e.g. out of range) or that the resulting dynamic range is too sensitive to the building block specifications. Then, a “high” cost is assigned (e.g. 10^7) to such solutions. The rejection of solutions that are too sensitive to the building block specifications ensures the robustness of the obtained parameter vector. This sensitivity information is returned by the simulator together with the dynamic range. As an example, the first solution point marked in figure 16.8 will be accepted as a stable design point, whereas the second solutions point is much too sensitive to parameter variations. The arrows indicate the (user-defined) span (or deviation) of the parameter that is used for calculating the sensitivity. The horizontal lines indicate the (user-defined) allowed variation in SNR. Given the nonlinear behavior of the simulation result, it is important to take margins for the flat region in figure 16.8.

Termination criteria

Every population member of each new generation being calculated by the optimization algorithm is evaluated by calculating its cost (*fitness*) according to equa-

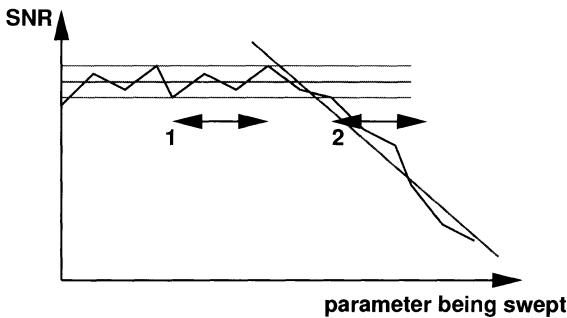


Figure 16.8: Sensitivity checking to ensure robust design points.

tion (16.6). A number of stop criteria have been incorporated into the algorithm in order to terminate the optimization under *reasonable circumstances*.

Power model

A simple but effective model has been implemented as relative power. A first consideration we have made is the fact that the largest part of the power consumption of the modulator is determined by the operational amplifiers. This is reflected in the following equation:

$$P \sim I_{BIAS_OTA} \sim \frac{g_m V_{GST}}{2} \quad (16.7)$$

On the other hand, we have:

$$GBW_{OTA} = \frac{g_m}{2\pi C_{eq}} \quad (16.8)$$

and thus:

$$P \sim C_{eq} GBW V_{GST} \quad (16.9)$$

When designing for minimum OTA power consumption, one will try to maximize g_m by choosing the gate overdrive voltage $V_{GST} \equiv V_{GS} - V_T$ as low as possible. The following expression is then taken as a relative power estimator:

$$P \sim GBW C_{eq} \quad (16.10)$$

where C_{eq} is the equivalent output load capacitance. In our estimator, the order of the modulator topology under test is taken into account. This includes also a scaling coefficient between the different stages which is also done in a practical implementation and thus reduces the power consumption of successive stages. In addition, a penalty coefficient was added for each OTA's power figure as a function

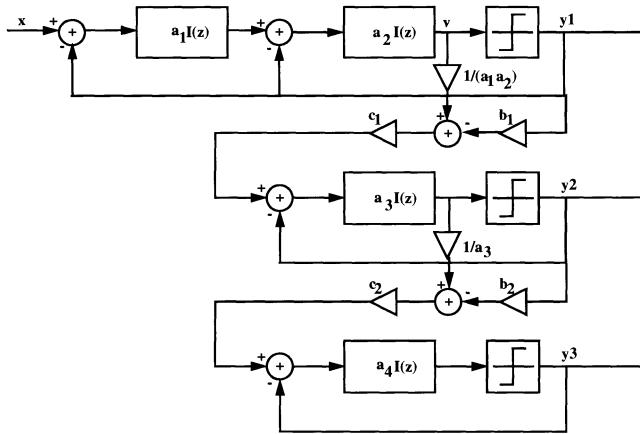


Figure 16.9: The fourth-order 2–1–1 cascade $\Delta\Sigma$ modulator used in [13].

of the gain requirement. Large gains will require e.g. gain-boosting stages [13] and thus consume more power.

16.4.1.2 Experimental results

In order to prove the validity of the high-level synthesis methodology, the following ADSL specifications were entered as input to the tool:

- peak SNR : 79 dB (after brick-wall filtering)
- signal bandwidth : 1.1 MHz
- V_{ref} : 0.9 V

These are similar as the ones used in the manual design of [13] where a fourth-order 2–1–1 cascade $\Delta\Sigma$ modulator was used to achieve these specifications. This architecture is shown in figure 16.9. The oversampling ratio used in [13] was 24 resulting in a sampling frequency of 52.8 MHz. Table 16.1 presents a comparison of the results published in [13] with the ones that were obtained by the synthesis tool.

These results are very close to the ones in [13] with exactly the same topology and oversampling ratio chosen by the synthesis tool. Of course, large design margins on the block specification are taken by designers in a real design, also to compensate for process variations, whereas the tool returns minimum specifications. Furthermore, it was found that in this optimization problem there are many solutions that give *fitness* values that are very close to each other. Since the genetic algorithm starts

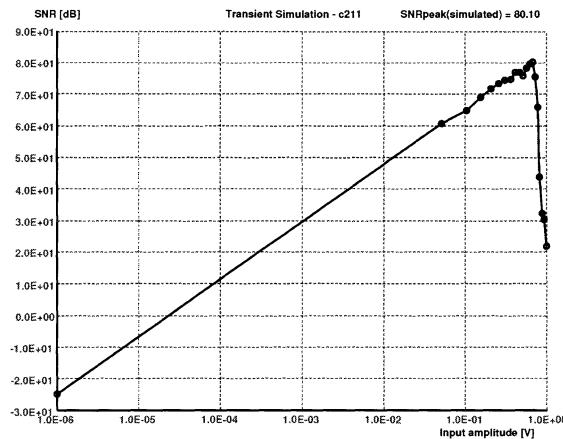


Figure 16.10: The simulation of SNR versus normalized input amplitude as verification.

from a random starting point, the final solution vector might be somewhat different, without being less *optimal*. In other words, because a multi-objective optimization is performed, the resulting solution is Pareto-optimal.

After the optimization, which took less than 8 minutes (including constant updating on the screen and logging to file) on an Intel PIII 700 MHz PC, the user gets a full report on the optimization history and an overview of the optimized variables. Also, the best solution is written back to the simulator data structures, so that a

Table 16.1: Comparison of the synthesized results for state-of-the-art specifications with published results of a working design.

building block specs	constraint	published [13]	synthesized
topology	—	cascade 2–1–1	cascade 2–1–1
oversampling ratio	—	24	24
OTA gain	>	60 dB	52 dB
OTA GBW	>	160 MHz	222 MHz
OTA output swing	>	1.8 V	1.86 V
switch on resistance	<	215 Ω	217 Ω
comparator offset	<	100 mV	130 mV
comparator hysteresis	<	40 mV	19 mV

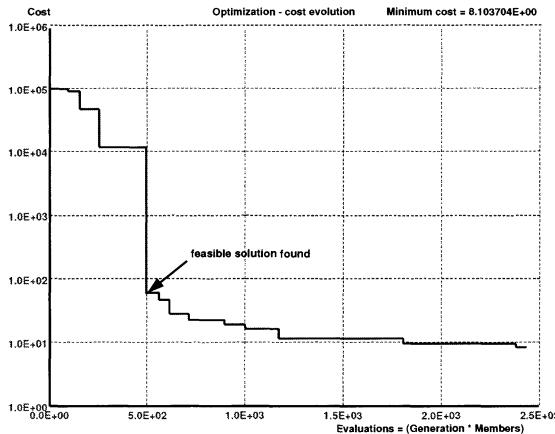


Figure 16.11: The minimum cost evolution during optimization.

verification can be done. The verified SNR versus input amplitude curve of the above example is shown in figure 16.10.

Figure 16.11 shows a graphical representation of the minimum cost evolution during the synthesis run. It is clearly visible that the algorithm quickly finds a feasible solution (corresponding to the largest cost decrease) and then further optimizes for minimum power consumption.

16.4.2 Systematic design of a Nyquist-rate D/A converter

For high-speed, high-accuracy D/A converters, a segmented current-steering topology is usually chosen as it is intrinsically faster and more linear than other architectures [14, 15]. The conceptual block diagram of this type of D/A converter is depicted in figure 16.12: the l least significant bits are implemented in a binary way while the m most significant bits steer a unary current source array. The general specification list for a current-steering D/A converter is given in figure 16.13. The specification values included will be used as example design throughout this entire section. The specifications can be divided into four categories: *static*, *dynamic*, *environmental* and *optimization* specifications. In the case of a D/A converter the *static* parameters include resolution (i.e. number of bits), integral non-linearity (INL), differential non-linearity (DNL) and yield. The *dynamic* parameters include settling time, glitch energy, spurious-free dynamic range (SFDR) and sampling frequency. The *environmental* parameters include the power supply, the digital levels, the output load and the input/output range. The power consumption and

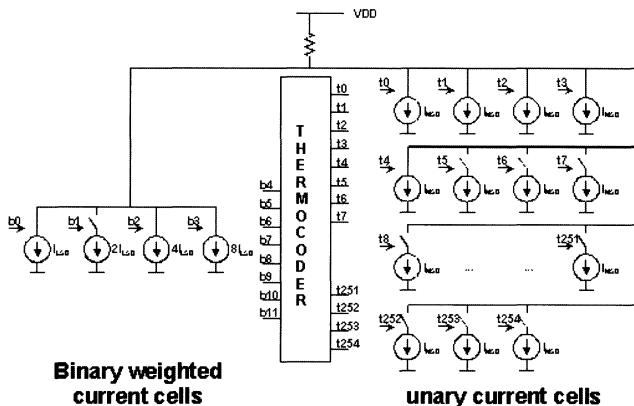


Figure 16.12: Principle block diagram of a 12-bit segmented current-steering D/A converter. The l ($=4$) least significant bits steer the binary-weighted current sources directly. The m ($=8$) most significant bits are fed into the thermometer decoder (thermocoder), which steers the unary current source array.

	Specification	Unit	Value	Design
Static	Number of bits (n)	-	14	14
	INL	LSB	0.5	0.3
	DNL	LSB	0.5	0.2
	Parametric Yield	%	99.9	NM*
Dynamic	Glitch energy	pV.s	1.0	NM*
	Settling time (10-90%)	Ns	10	0.9
	SFDR @ 500KHz	dB	80	84
	Sample frequency	MHz	100	150
Environmental	Output range (V_{swing})	V	0.5	0.5
	R_{Load}	Ω	25	25
	Digital levels	-	CMOS	CMOS
	Power Supply	V	2.7	2.7
	Technology	-	0.5 μ 1P3M	0.5 μ 1P3M
Optimization	Power	mW	Min. (300)	300
	Area	mm^2	Min. (10)	3.2 x 4.1

Figure 16.13: Specification list for a current-steering D/A converter with specification values and measured values of a fabricated design [16]. (*NM is not measured).

area are the optimization targets and need to be minimized for a given technology. This specification list serves as input for the design process to be explained next.

16.4.2.1 Architectural design

In the architectural design phase mainly two important architectural decisions have to be taken. The first one is the decision about how many bits (l) are implemented using binary-weighted current sources and how many (m) using unary-weighted sources. The second important architectural design decision of a current-steering D/A converter is the switching scheme. The switching scheme has two components. A unary current source can be composed of one or more parallel units spread out over the current source array to average out any systematic gradients, which can be necessary for high-accuracy applications. The second parameter of the switching scheme is the switching sequence, i.e. the sequence in which the different current sources are turned on for increasing input code, in relation to their geometrical position on the layout of the current source array. It can be shown that the remaining spatial errors are not accumulating when the current sources are switched on in some optimal way [16]. It is important that any design methodology offers sufficient flexibility to make these architectural explorations possible.

To decide on the amount of segmentation (l, m), both static and dynamic performance criteria are taken into account. The static behavior of a D/A converter is specified in terms of INL and DNL. A distinction has to be made between random errors and systematic errors. The random errors are solely determined by mismatches. The systematic errors are caused by process, temperature and electrical gradients. In optimally designed D/A converters the systematic errors should be minimized by design and careful layout, and hence the INL and DNL should be limited by random errors (i.e. mismatch) only. A small safety margin (10 % of INL) is reserved to allow for systematic contributions. Since the random errors are caused by mismatches, the maximum acceptable random error on the LSB current can be calculated from yield simulations [17]. For example, for the 14-bit D/A converter specified in figure 16.13, and to achieve a targeted yield that $INL < 0.5LSB$ for 99.9 % of all fabricated IC's, the relative standard deviation of current matching for the unit current cell (1 LSB) $\sigma(I)/I$ has to be smaller than 0.1 %. Then, an estimate for the active area of the current source transistor can be calculated based on the mismatch model [17]:

$$W * L = \frac{1}{2\sigma^2(I)} \cdot \left(A_\beta^2 + \frac{4A_{VT}^2}{(V_{GS} - V_T)^2} \right) \quad (16.11)$$

where $\sigma(I)/I$ is the unit current source standard deviation and A_β, A_{VT} are technology constants that characterize the MOS transistor mismatch. Typically, for

minimal area the $(V_{GS} - V_T)$ is maximized. The total current source array area $area_{cur_{src,est}}$ can then be estimated:

$$area_{cur_{src,est}} \approx 2 \cdot (W \cdot L)_{lowerbound} \cdot f_{routing} \quad (16.12)$$

where $f_{routing}$ is the routing overhead factor. The static performance places a strict minimum on the area of the current source array.

The dynamic behavior of a D/A converter is usually specified in terms of admissible glitch energy. This specification is mainly determined by 1) the number of bits implemented in a unary/binary way, and 2) the way the current sources are synchronized when switched on/off. The largest glitch will occur when switching off all binary implemented bits and switching on the first unary current source. This implies that the decision on the number of bits l to be implemented in a binary way and the number of bits m to be implemented in a unary way determines the worst-case glitch. The lowest possible glitch energy is obtained when a full unary implementation is chosen [18]. This would however result in a large area increase because of the complex binary-to-thermometer encoder that is needed to convert the digital binary input codes into thermometer code to steer the unary current cells. The area of the current source array is, however, fixed by equation (16.12) as is the area of the array of the switches/latches. On the other hand, the area of the thermocoder increases ($\sim 2m$), as does the size of the routing busses connecting the three modules, if the number of unary bits (m) is increased. Figure 16.14 shows that in today's technologies an optimal number of unary implemented bits is around 8 bits for a total of 14 bits of resolution, otherwise the area grows unacceptably large. This choice will ultimately limit the dynamic performance of the D/A converter.

The next decision to be taken is the switching scheme. Due to the size of the array, there will be large gradients (electrical, thermal, etc.) from left to right, and from top to bottom, over the array of current sources. Hence, this will result in systematic errors and therefore in INL and DNL nonlinearities. This can be reduced by carefully placing the current sources across the array and by switching them on in a specific sequence. To compensate linear gradients and to reduce nonlinear gradients, a unary current source can be split into multiple smaller current sources, spread over the array and connected in parallel. At least 4 subunits are required for 12-bit linearity [19] and 16 subunits for 14-bit linearity [16]. To reduce the accumulation of the residual systematic errors when running through the digital input codes, the sequence in which the different current sources have to be turned on in relation to their position on the array layout, has to be carefully chosen. An optimized switching sequence that randomizes rather than accumulates the residual errors, has been determined using a branch and bound search algorithm [16]. The result is shown in figure 16.15. On the left a manually derived switching sequence

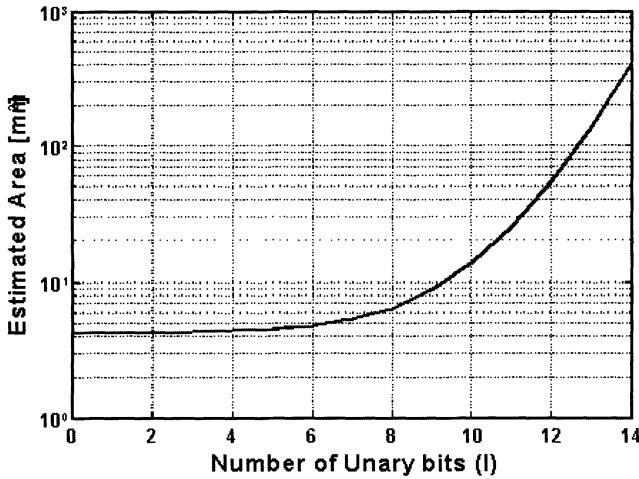


Figure 16.14: Estimated area of the D/A converter as a function of the number of unary bits.

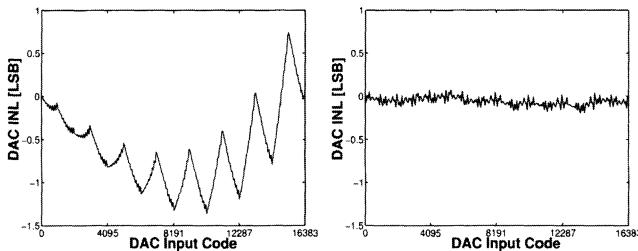


Figure 16.15: Prediction of the INL of a 14-bit D/A converter obtained with two different switching sequences. On the left a manually derived (suboptimal) switching sequence; on the right the computer-generated optimized switching sequence. The same error profile (extracted from a test chip) has been applied in both cases.

has been simulated which was sufficient for the 12-bit converter in [20] but clearly is insufficient for 14 bits; on the right with the same spatial errors the optimized switching scheme from [16] results in an INL of 0.2 LSB instead of 2.1 LSB.

16.4.2.2 Circuit-level design

Now that the architecture is fully determined, in the next phase the devices sizes and biasing of all circuits in the converter have to be determined. These circuits

include the current sources, the switches, the latches, but also the digital circuitry. For the sizing of the analog circuits, again the two performance constraints – static and dynamic – are taken into account. The active area of the unit current source is calculated from mismatch constraints (see equation (16.11)). A high biasing voltage ($V_{GS} - V_T$) is preferred for mismatch reasons. The upper limit for the biasing voltage is determined by the output swing (switching transistors need to be in the saturation region) and by the power supply. Hence, using equation (16.11) and knowing the output range and load, $W_{cur_{src}}$ and $L_{cur_{src}}$ can be calculated for some proper choice of ($V_{GS} - V_T$). To avoid distortion due to the nonlinearity of the output impedance when a different number of current sources are switched to the output, a cascode transistor is typically inserted in the current source to make the impedance sufficiently high. The calculations, which result in the sizing of the current source transistors ($M_{cur_{src}}$ and M_{casc}), can be implemented in a MATLAB® script. For the 14-bit converter of figure 16.12 this script gives a $W/L = 1.1\mu m/104.5\mu m$ in a standard $0.5\mu m$ CMOS technology.

In order not to deteriorate the dynamic performance, the following factors are taken into account in the circuit-level synthesis [21]: 1) synchronize the control signals of the switching transistors, 2) reduce the voltage fluctuation in the drains of the current sources during switching, 3) carefully switch the current source transistor on/off. The synchronization of the control signals is achieved by adding a latch immediately in front of the switching transistors. The voltage fluctuation at the drain changes the current from the current source because of the finite output impedance of the current source transistor. The problem can be solved by using a large channel length for the current source transistor, and tuning the crossing point of the switching control signals such that both switches are never switched off simultaneously [20]. Using a SPICE simulator within an optimization loop, the latch and the switches are sized automatically, taking the above crossing point and speed considerations as constraints in the optimization process.

Finally, as the architectural parameters (l, m) and the latch transistor sizes are now known, the digital thermometer decoder can be synthesized. The remaining l LSBs are delayed by an equalizer block to have the same overall delay. The full decoder is synthesized from a VHDL description using a logic synthesis tool, and enforcing the required timing under the known load constraints.

16.4.2.3 Layout generation

Like most Nyquist-rate converters, current-steering D/A converters are a typical example of layout-driven analog design. The sized schematic alone does not constitute an operational converter. An important part of the performance is determined by the handling of layout-induced parasitics and error components (i.e. systematic

errors). All classical countermeasures for digital to analog coupling (guard rings, shielding, separate supplies, etc.) and standard matching guidelines (equal orientation, dummies, etc.) are typically applied and will not be discussed further. We will concentrate here on the extra required layout measures and the Mondriaan layout tool [22], specially developed for such regular layout structures.

In a first step of generating the layout of a converter, the global floorplan is determined. This indicates the relative position of the major modules in the layout, and the global interconnection strategy. For the example 14-bit D/A converter it was decided to have the digital decoder, the switch/latch array and the current source array as three modules next to each other in the layout. To make a square or near-square overall chip layout, this constrains the aspect ratio of the three modules and serves as input for their individual layout generation. Secondly, at the chip level the connections between the three modules are extremely important: to avoid waisted area e.g. due to river routing, the modules should be connected by abutment and therefore a fixed wiring pitch must be chosen to route the busses across the different modules.

The layouts of the current source array and swatch array are generated next. The sizes of the unit current source for the example of figure 16.13 have been determined above. From this the sizes of all other weighted current sources and the unary current source are derived. For optimal matching the current source array must be built up from identical basic units. This basic unit is laid out manually. It contains the layout of a current source and optionally a cascode transistor. The placement of the basic units in the array is known from the switching scheme. The current source array is then generated automatically with the Mondriaan tool [22]. Mondriaan is a dedicated analog layout synthesis tool targeted to the automated layout generation of highly regular array-type analog modules. It takes the complex switching scheme and the wiring pitch as inputs, and performs a floorplanning, symbolic routing and technology mapping to generate the physical layout of the entire current source array. Run times are in the order of a minute. The basic swatch cell is also laid out manually. The placement and routing of the entire switch/latch array is then again done automatically with the Mondriaan tool [22]. Inputs to the tool are the pin list of the current source array, the netlist, the wiring pitch, the layout of one individual swatch cell, and some technology data.

Finally, the layout of the digital thermometer decoder is generated using a commercial digital standard cell place and route tool, using the pin list obtained from the swatch array layout as input. The modules are then stacked on top of each other. The bus generators of the Mondriaan tool [22] are used to generate the connections between the three modules (full decoder, swatch array and current source array). Trees are used to collect the output signals and distribute the clocking signal

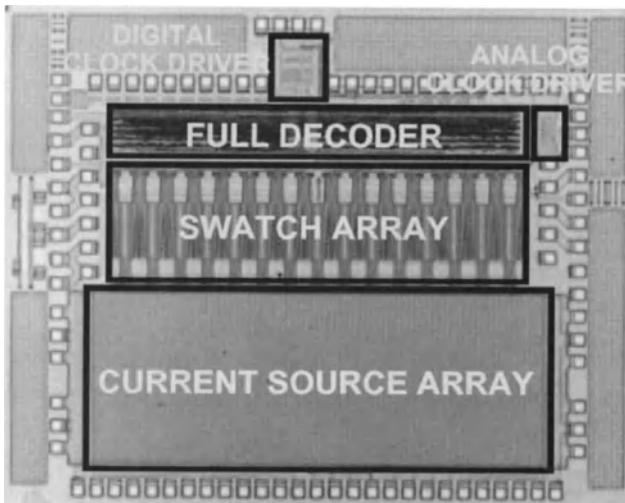


Figure 16.16: Layout of a 14-bit segmented current-steering D/A converter.

from the clock driver to the swatch array to have equal delay. The bonding pads are placed and manually connected to all the external pins of the D/A converter. For the example 14-bit current-steering D/A converter design of figure 16.12, this results in the layout shown on the microphotograph of figure 16.16 [16]. This design was fabricated in a $0.5 \mu\text{m}$ CMOS technology. The measurement results are included in the last column of figure 16.13 [16]. The update rate of this converter is 150 MSamples/s. The chip has a measured INL of 0.3 LSB and a DNL of 0.2 LSB. This proves that the approach of using an optimized switching scheme is absolutely needed for 14-bit accuracy. A SFDR of 61 dB was measured for a full-scale input signal at 5 MHz.

16.4.2.4 Design speed-up

This concludes the complete sizing and layout process of the D/A converter. The methodology has been applied to three different current-steering D/A converters, that have all been fabricated and measured [23]. The time spent on the different steps in the design methodology has been counted. During the course of the first design no tools were available, and all steps were performed manually. During the second design certain MATLAB® scripts were used and the Mondriaan tool was under development. In the third design all scripts and the Mondriaan tool were fully ready. Using the MATLAB® sizing scripts, the design time for sizing could be

reduced from 4 weeks down to 1 week. Using the Mondriaan tool the layout time of the analog arrays was reduced from 3 weeks down to 1 day. No reduction on the digital design times was obtained, since the same commercial tools were used in all three designs. Similarly, in all three designs, the final verification (DRC and LVS, parasitics extraction, backannotation and final verification with HSPICE) required 2 weeks – a time that we did not succeed to cut down.

In summary, by using the tool-supported systematic design methodology for analog building blocks, the overall design time was reduced from 11 weeks for the first D/A converter design down to 4 weeks of total personneffort for the third D/A converter design. This is a reduction of the total design time by a factor of 2.75, which experimentally indicates the boost in productivity obtained in the design of analog macrocells by using a tool-supported systematic design methodology. And by automating more and more of the design steps, such as for instance by automating the entire high-level architectural design of a $\Delta\Sigma$ modulator by means of the DAISY tool [10], or the entire circuit-level design of a converter's building blocks by means of the AMGIE tool [2], the productivity gain can be increased even further. Note, however, that full automation of the design of analog macrocells is only useful and meaningful for those types of blocks that show a high level of topological reuse, such as $\Delta\Sigma$ converters and pipelined converters, whereas other types of blocks show more topological variations and therefore require a more flexible approach. Yet, as was shown above for the example of current-steering D/A converters, even for these types of designs a tool-supported systematic design methodology can drastically reduce the design time and this does not come at the expense of performance. Much like Gaudi's beautiful work looks virtuoso but is actually based on careful design and engineering.

16.5 EXTRACTED BEHAVIORAL MODEL FOR BOTTOM-UP VERIFICATION

After the design and the layout of the converter are completed and verified, the last phase in the design flow of embedded data converters (see figure 16.1) is to generate a behavioral model in which the model parameters are extracted from the designed circuit. The resulting model is used for final verification of the system in which the data converter is used as an embedded functional block. Note that this model can be more accurate than the generic behavioral model that is used during the top-down system exploration phase, since now the exact implementation of the functional block is known. This phase is again illustrated with the extracted behavioral model of both an analog-to-digital and a digital-to-analog Nyquist-rate data converter.

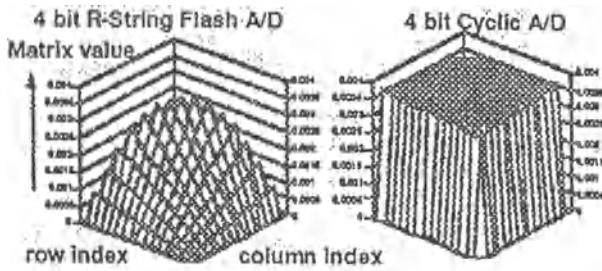


Figure 16.17: Variance–covariance matrix of the transition points of the static transfer curve of a 4-bit resistor–string flash A/D converter (left) and a 4-bit algorithmic A/D converter (right).

It was described in equation (16.2) how the static input–output relationship of a Nyquist–rate data converter can be fully described by the transition points in the transfer curve of the converter [8]. The deterministic μ_t and the statistical Σ_t parts of these transition points can be extracted from the actual design of the circuit, the latter one using mismatch simulations. To reduce the size and since the rank of the variance–covariance matrix Σ_t can be much smaller than 2^n , it can be reduced by means of singular value decomposition into a smaller matrix Σ_{ct} . Hence, any statistical sample of the designed converter is fully characterized by the following equation [8]:

$$t = \mu_t + U_t c_t, c_t \sim \text{normal}(0, \Sigma_{ct}) \quad (16.13)$$

where:

$$\Sigma_t = U_t \Sigma_{ct} U_t^T \quad r = \text{rank}(\Sigma_t) \quad (16.14)$$

Figure 16.17 illustrates the extracted Σ_t matrix of both a 4-bit resistor–string flash A/D converter and a 4-bit algorithmic A/D converter [8]. It is clearly visible that the mismatch errors in the flash converter are independent and hence the Σ_t matrix is full rank, whereas the cyclic reuse of hardware in the algorithmic converter results in a rank of only 3 for the Σ_t matrix. This means that there are only 3 independent error vectors in the entire converter, and that any fabricated sample of this converter will have an INL and DNL that is a linear combination of these 3 vectors (with coefficients depending on the actual mismatches on that fabricated sample).

For the dynamic (transient) behavior of a D/A converter, a model for the glitch was presented in equation (16.3) [9]. In the bottom–up extraction phase, this glitch model can be extended to reflect better the actual design. Firstly, a separate damped sine can be used for switching on ($A_{gl, on}$ and $t_{gl, on}$) respectively switching off ($A_{gl, off}$ and $t_{gl, off}$) a current source. The number of current sources that are switched on or

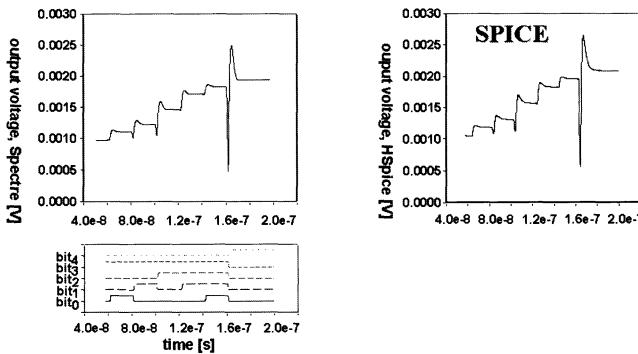


Figure 16.18: Comparison between the extracted behavioral model and the numerical device–level HSPICE simulation of the time response of a current–steering D/A converter.

off can readily be computed from the chosen topology (i.e. the choice of the number of bits l that steer the binary–weighted current source array). Switching on or off current sources is thus modeled separately and then combined at the output node. Secondly, the amplitude and the time constant of the damped sine of the left–hand side and the right–hand side (see figure 16.3) are controlled separately ($A_{gl, left}$ and $t_{gl, left}$ and $A_{gl, right}$ and $t_{gl, right}$). This results in 8 parameters for the final model, that can easily be extracted from SPICE circuit simulations by switching one bit on and off. The comparison between a circuit–level HSPICE simulation and the response of the extracted behavioral model for the same digital input is depicted in figure 16.18. The extracted model has an accuracy better than 99 percent, or an error less than 1 percent. The HSPICE simulation required 4:37 minutes of CPU time on a HP 712/100 whereas the behavioral model only took 12 seconds. This implies a simulation speed–up of 874 times, which is needed for system–level verifications.

16.6 CONCLUSIONS

A systematic design methodology has been presented for CMOS data converters that are used as embedded macrocells in integrated VLSI systems. The methodology covers the complete design flow starting from the specification phase of the converter as macrocell in a larger system down to a fully synthesized and laid out implementation, followed by a verification of the entire system with a behavioral model extracted from the actual designed converter. First, a generic behavioral model, that includes the most important nonidealities in a parameterized way, is used for system–level exploration to define the converter’s individual specifications

within the system. The architecture of the converter and the sizes of the devices are then determined using a two-level performance-driven design methodology. At the converter architectural level the architectural decisions (the exact converter topology and the subblock requirements) are determined. This can be automated by combining an optimization algorithm with a fast behavioral simulation of the converter. The derived subblock specifications then serve as input for the circuit sizing (optimization) of the individual subblocks. This is then followed by the layout generation phase. This procedure has been illustrated with the automatic high-level synthesis of a $\Delta\Sigma$ modulator, using the DAISY tool, as well as with the entire systematic design of a 14-bit Nyquist-rate current-steering D/A converter. The latter has been supported by both commercially available and newly developed software tools (such as the Mondriaan layout tool that was specially developed for the generation of the layout of regular array-type analog structures with complex connectivity patterns, such as encountered in Nyquist-rate data converters). Finally in the design methodology, a detailed behavioral model is extracted from the completed design, combining in accurate detail the major nonidealities such as static (e.g. INL, DNL) and dynamic (e.g. glitch, SFDR) behavior. This has also been illustrated with several modeling examples. The experimental results have demonstrated how the systematic design methodology and the supporting tools drastically reduce the total design time for embedded data converters with a factor 2.5 and above (depending on the level of automation), thereby significantly increasing analog design productivity, without sacrificing performance.

ACKNOWLEDGEMENTS

This work has been supported in part by ESA-ESTEC under the GSTP program, and by the SYSCONV project under the Mixed-Signal Initiative of the European Union.

REFERENCES

- [1] Georges Gielen and Rob Rutenbar, "Computer-Aided Design of Analog and Mixed-Signal Integrated Circuits," *Proceedings of the IEEE*, vol. 88, no. 12, pp. 1825–1854, December 2000.
- [2] Geert Van der Plas, Geert Debyser, Francky Leyn, Koen Lampaert, Jan Vandebussche, Georges Gielen, Willy Sansen, Petar Veselinovic, and Domine Leenaerts, "AMGIE – A Synthesis Environment for CMOS Analog Integrated Circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 20, no. 9, pp. 1037–1058, September 2001.
- [3] R. Phelps, M. Krasnicki, R. Rutenbar, L. R. Carley, and J. Hellums, "ANACONDA: Simulation-Based Synthesis of Analog Circuits Via Stochastic Pat-

- tern Search,” *IEEE Transactions on Computer–Aided Design of Integrated Circuits and Systems*, vol. 19, no. 6, pp. 703–717, June 2000.
- [4] F. Medeiro, B. Perez-Verdu, A. Rodriguez-Vazquez, and J. Huertas, “A Vertically Integrated Tool for Automated Design of $\Delta\Sigma$ modulators,” *IEEE Journal of Solid-State Circuits*, vol. 30, no. 7, pp. 762–772, 1995.
 - [5] M. del Mar Hershenson, S. Boyd, and T. Lee, “Optimal Design of a CMOS Op-Amp Via Geometric Programming,” *IEEE Transactions on Computer–Aided Design of Integrated Circuits and Systems*, vol. 20, no. 1, pp. 1–21, January 2001.
 - [6] G. Gielen and J. da Franca, “CAD Tools for Data Converter Design: An Overview,” *IEEE Transactions on Circuits and Systems II*, vol. 43, no. 2, pp. 77–89, February 1996.
 - [7] P. Kinget and M. Steyaert, “Impact of Transistor Mismatch on the Speed–Accuracy–Power Trade–Off of Analog CMOS Circuits,” in *Proceedings Custom Integrated Circuits Conference (CICC)*, 1996, pp. 333 – 336.
 - [8] E. Liu, A. Sangiovanni-Vincentelli, G. Gielen, and P. Gray, “A Behavioral Representation for Nyquist-rate A/D Converters,” in *Proceedings IEEE/ACM International Conference on Computer–Aided Design (ICCAD)*, 1991, pp. 386 – 389.
 - [9] J. Vandenbussche et al., “Behavioral Model for D/A Converters as VSI Virtual Components,” in *Proceedings Custom Integrated Circuits Conference (CICC)*, May 1998, pp. 473 – 477.
 - [10] K. Francken, P. Vancorenland, and G. Gielen, “DAISY: A Simulation–Based High–Level Synthesis Tool for $\Delta\Sigma$ Modulators,” in *Proceedings IEEE/ACM International Conference on Computer–Aided Design (ICCAD)*, November 5 – 9 2000, pp. 188 – 192.
 - [11] K. Francken, M. Vogels, and G. Gielen, “Dedicated System–Level Simulation of $\Delta\Sigma$ Modulators,” in *Proceedings Custom Integrated Circuits Conference (CICC)*, May 6 – 9 2001, pp. 349 – 352.
 - [12] R. Storn, “On the Usage of Differential Evolution for Function Optimization,” in *NAFIPS*, 1996, pp. 519–523.
 - [13] Y. Geerts, A. Marques, M. Steyaert, and W. Sansen, “A 3.3 V 15-bit Delta–Sigma ADC with a Signal Bandwidth of 1.1 MHz for ADSL-applications,” *IEEE Journal of Solid-State Circuits*, vol. 34, no. 7, pp. 927–936, February 7–9 1999.
 - [14] Behzad Razavi, *Principles of Data Conversion System Design*, IEEE Press, 1995.
 - [15] R. van de Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters*, Kluwer Academic Publishers, 1993.
 - [16] Geert Van der Plas, Jan Vandenbussche, Willy Sansen, Michiel Steyaert, and Georges Gielen, “A 14-bit Intrinsic Accuracy Q^2 Random Walk High-Speed

- CMOS DAC,” *IEEE Journal of Solid-State Circuits*, vol. 34, no. 12, pp. 1708–1718, December 1999.
- [17] J. Bastos, M. Steyaert, and W. Sansen, “A High Yield 12-bit 250 MS/s CMOS D/A Converter,” in *Proceedings Custom Integrated Circuits Conference (CICC)*, May 1996, pp. 431 – 434.
 - [18] C-H. Lin and K. Bult, “A 10b 500 MSamples/s CMOS DAC in $0.6 \mu\text{m}^2$,” *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 1948–1958, December 1998.
 - [19] J. Bastos, A. Marques, M. Steyaert, and W. Sansen, “A 12-bit Intrinsic Accuracy High-Speed CMOS DAC,” *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 1959–1969, December 1998.
 - [20] A. Van den Bosch et al., “A 12 bit 200 MHz Low Glitch CMOS D/A Converter,” in *Proceedings Custom Integrated Circuits Conference (CICC)*, May 1998, pp. 249 – 252.
 - [21] T. Wu et al., “A Low Glitch 10-bit 75-MHz CMOS Video D/A Converter,” *IEEE Journal of Solid-State Circuits*, vol. 30, no. 1, pp. 68–72, January 1995.
 - [22] G. Van der Plas, J. Vandebussche, G. Gielen, and W. Sansen, “A Layout Synthesis Methodology for Array-Type Analog Blocks,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 21, no. 6, pp. 645–661, June 2002.
 - [23] J. Vandebussche et al., “Systematic Design of High-Accuracy Current-Steering D/A Converter Macrocells for Integrated VLSI Systems,” *IEEE Transactions on Circuits and Systems II*, vol. 48, no. 3, pp. 300–309, March 2001.