

Brought to You by



Like the book? Buy it!

Xuejun Zhang
Lawrence E. Larson
Peter M. Asbeck



Design of Linear RF Outphasing Power Amplifiers

Design of Linear RF Outphasing Power Amplifiers

For a listing of recent titles in the *Artech House Microwave Library*,
turn to the back of this book.

Design of Linear RF Outphasing Power Amplifiers

Xuejun Zhang
Lawrence E. Larson
Peter M. Asbeck



Artech House
Boston • London
www.artechhouse.com

Library of Congress Cataloging-in-Publication Data

A catalog for this book is available from the U.S. Library of Congress.

British Library Cataloguing in Publication Data

Zhang, Xuejun

Design of linear RF outphasing power amplifiers.—(Artech House microwave library)

1. Power amplifiers—Design 2. Amplifiers, Radio frequency—Design

I. Title II. Larson, Lawrence E. III. Asbeck, Peter

621.3'8412

ISBN 1-58053-374-4

Cover design by Igor Valdman

© 2003 ARTECH HOUSE, INC.

685 Canton Street

Norwood, MA 02062

All rights reserved. Printed and bound in the United States of America. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system without permission in writing from the publisher.

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Artech House cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

International Standard Book Number: 1-58053-374-4

A Library of Congress Catalog Card Number is available from the U.S. Library of Congress.

10 9 8 7 6 5 4 3 2 1

To my wife Huijuan—always part of what I do
— X. Zhang

To the memory of my late father, Clarence Edward Larson
— L. E. Larson

To my wife, Marcia, and my children, Alan and Lynn
— P. M. Asbeck

Contents

	Preface	<i>xi</i>
1	Introduction	1
1.1	The Role of Power Amplifiers in Wireless Communication Systems	1
1.2	Characterization of Power Amplifiers for Wireless Communications	4
1.2.1	Power Amplifier Waveform Quality Measurements	6
1.2.2	Power Efficiency Measurements	19
1.3	Power Amplifier Linearization and Efficiency-Enhancement Techniques	22
1.4	Outphasing Microwave Power Amplifiers	27
1.4.1	Historical Perspectives on Outphasing Power Amplifiers	27
1.4.2	Introduction to the Theory of Outphasing Amplification	29
	References	32
2	Linearity Performance of Outphasing Power Amplifier Systems	35
2.1	Introduction	35
2.2	Digital Modulation Techniques	36
2.2.1	QPSK and Its Variations	36
2.2.2	QAM	41

2.3	Baseband Filtering of Digital Data	41
2.3.1	Raised Cosine Filter	45
2.3.2	Gaussian Filter	49
2.3.3	IS-95 Baseband Filter	50
2.4	Signal Component Separation for Outphasing Amplifiers	51
2.5	Path Imbalance and Its Effects on Linearity	61
2.5.1	Two-Tone Linearity Analysis of an Outphased Amplifier with Path Mismatch Effects	62
2.5.2	ACI Estimation with Gain and Phase Mismatch	64
2.6	Effect of Quadrature Modulator Errors on Linearity	67
2.6.1	Quadrature Modulator Error Minimization	69
2.6.2	Quadrature Modulator Error Effects on Outphasing Systems	72
2.7	SCS Quantization Error Effects on Outphasing Systems	75
2.7.1	Error Effects of Quantization of the Source Signal	75
2.7.2	Error Effects of Quantization of the Quadrature Signal	76
2.8	Linearity Effects of Reconstruction Filter and DSP Sampling Rate	79
2.9	Summary	81
	References	83
3	Path Mismatch Reduction Techniques for Outphasing Amplifiers	87
3.1	Introduction	87
3.2	Correction Schemes Based on Training Vectors	88

3.2.1	Baseband Preconditioning of Path Mismatch Errors	89
3.2.2	Foreground Calibration Algorithm of Path Mismatch Errors	91
3.3	Path Mismatch Error Correction Schemes Transparent to Data Transmission	101
3.3.1	Phase-Only Correction Approach	101
3.3.2	Simplex Search Algorithm Correction	102
3.3.3	Direct Search Algorithm	104
3.3.4	Background Calibration Algorithm	105
3.4	Mismatch Correction Scheme for Broadband Applications	112
3.5	VCO-Derived Synthesis	114
3.5.1	CALLUM	119
3.5.2	VLL	124
	References	126

4 Power-Combining and Efficiency-Enhancement Techniques 129

4.1	Introduction	129
4.2	Power-Combining Techniques for Outphasing Amplifiers	130
4.3	Amplifier Choices for Outphasing Systems	135
4.4	Outphasing Amplifier Design Using Class A, B, and C Amplifiers	137
4.5	Chireix Power-Combining Technique	142
4.6	Combiner Design for Switching-Mode (Class D and Class E) Amplifiers	145
4.6.1	Analysis of MOSFET-Based Class D Outphasing Amplifier with Lossless Combining	147
4.6.2	Simulation and Discussion of MOS-Based Class D with Lossless Combining	153

4.7	Application of Lossy Power Combiners to Outphasing Power Amplifiers	156
4.8	Probability Distribution of Output Power and Its Impact on Efficiency	159
4.9	Power Recycling in Outphasing Amplifiers	163
4.9.1	Analysis of the Power-Recycling Network for a Continuous-Wave Signal	164
4.9.2	Analysis and Discussion for Linear- Modulated Signals	176
4.9.3	Practical Implementations	184
	References	185
	Selected Bibliography	185
	Appendix 4A	186
4A.1	Available Power from the Hybrid Combiner	186
4A.2	Recycling Efficiency and VSWR for Arbitrary Diode Model	188
	About the Authors	191
	Index	193

Preface

The wireless communications revolution has been driven by a confluence of technological advances, including improvements in communications theory, very large-scale integration (VLSI) technology, and radio frequency (RF) microelectronics. The microwave power amplifier represents one of the major factors in the low-cost and low-power implementation of these systems, since they are required to deliver tens of watts of power (in the case of a base station) with exacting standards of linearity and direct current (dc) power efficiency. The microwave power amplifier in the handset must deliver—at most—a few watts of power but dissipate very little dc power and sell for only a few dollars in large quantities. Clearly, these requirements represent an enormous challenge for the design and implementation of the power amplifier circuit.

This challenge has been met through a variety of creative approaches over the years, and the linearization and efficiency enhancement of microwave amplifiers remains an area of active research. Generally speaking, linearization techniques can be classified under the categories of either *feedback* or *feed-forward* approaches. Both of these approaches have well-known advantages and disadvantages. Feedback approaches tend to exhibit stability problems, and feed-forward approaches suffer from matching limitations. Neither has achieved very widespread application in commercial communications systems.

A third technique for amplifier linearization is known as the *outphasing approach*. This approach also has a long history—dating back to the 1930s where it was first proposed for amplitude-modulated (AM) transmission—and uses power-combining techniques to create a *linear* output waveform from two *nonlinear* input sources. When combined properly, the resulting waveform is highly linear, but the input amplifiers that create the nonlinear input waveforms can be designed to be very efficient from a dc power perspective.

Although this approach has much to recommend it, it has not achieved widespread use in commercial applications. This is because the intrinsic

advantages are offset by some significant disadvantages; in particular, the matching requirements between the two amplifiers are very stringent, and a large fraction of the microwave power is typically wasted in the power-combining network.

However, there has been a flurry of activity recently in this area in an attempt to solve these problems, and the next generation of outphasing microwave power amplifiers is poised to solve these historical limitations and finally find its way into commercial applications. This book aims to provide the reader with the most up-to-date summary of recent advances in outphasing microwave power amplifier design and to present the historical background that led to their development.

Background Information and Guidelines

The book is organized as follows: Chapter 1 begins with a brief introduction of wireless communication standards and their imposed requirements on power amplifiers. It also introduces the concept of the outphasing power amplifier and discusses its historical advantages and limitations.

Chapter 2 discusses outphasing amplifier system linearity performance, first introducing the typical modulations and baseband filtering techniques. Then, Chapter 2 describes various schemes for implementing the signal component separation function. The major sources of imbalance, including the path imbalance, the quadrature modulator error, the signal component separator (SCS) quantization, the digital signal processing (DSP) sampling rate and reconstruction filtering are successively investigated and translated into the system adjacent channel power regrowth (ACPR) requirements.

Chapter 3 reviews several correction approaches to compensate the imbalance between the two amplifier paths and improve the linearity performance of the outphasing system. These include schemes based on training vectors and those that can operate in background. For broadband application, channel equalization must be used to balance the responses of two branches for the entire band. Chapter 3 also discusses unique methods based on voltage controlled oscillator (VCO)-derived synthesis. These alternative approaches to achieve signal component separation can automatically correct the path imbalance; this makes them particularly suitable for low-power and low-cost integrated circuit implementation.

Chapter 4 concludes the book by addressing the efficiency-linearity trade-off in the power combiners of the two amplifiers. Traditionally, much of the advantage of the outphasing approach is “thrown away” in the power

combining network. The chapter first discusses the limitations of the traditional approaches and then presents improved techniques. These techniques fall into the categories of *reactive-combining* and *power-recycling* approaches. Reactive-combining approaches minimize the dc power dissipation by changing the load impedance to be more highly reactive at low output powers. Power-recycling approaches reduce the wasted power by converting the out-of-phase power back to dc and returning it to the power supply.

Acknowledgments

No book of this scope could have been completed without the dedicated assistance of many of our colleagues and friends. The authors would like to acknowledge valuable discussions with Dr. Steve Cripps, Dr. Fredrick Raab, Dr. Thomas Hornak, Dr. Bo Shi, Mr. John Sevic, and Professor Ian Galton on various aspects of outphasing power amplifier design. The authors would especially like to acknowledge Dr. Lars Sundstrom of Ericsson for his pioneering work on outphasing amplifier design and many valuable discussions on this topic. In addition, the administrative support of Ms. Michell Parks, Arline Allen, and Mr. James Thomas of the University of California, San Diego (UCSD) is greatly appreciated.

The financial support of the UCSD Center for Wireless Communications and Dr. John Lavery of the Army Research Office and multiuniversity research initiative (MURI) program, Applications of Nonlinear Dynamics and Chaos to Digital Communications, is also gratefully acknowledged.

1

Introduction

1.1 The Role of Power Amplifiers in Wireless Communication Systems

The tremendous growth of the wireless market, coupled with the fierce competition of the last few years, has spurred unprecedented interest in the performance of low-cost and physically compact RF power amplifiers. Concern for performance was motivated by the significant impact the power amplifier has on the talk time of the mobile station, which can represent a substantial marketing advantage if it is well-optimized. In addition, reducing the power requirements on the base station can have a significant impact on cost, and the power amplifier has a significant role to play there as well.

Classical power amplifier design techniques frequently ignore signal characteristics, instead focusing on transistor performance and circuit design techniques. However, with the emergence of digital wireless communication systems, an understanding of digital modulation theory is also required to fully characterize the power amplifier performance with the modulated carrier for the optimum linearity-efficiency trade-off. This section briefly reviews the first generation (1G) and second generation (2G) analog and digital wireless communication systems, with emphasis on the influence of wireless standards on power amplifier performance requirements. Details on wireless theory and communication theory can be found in [1–5].

Cellular wireless systems and standards have been in the process of evolution for decades throughout the world. The advanced mobile phone system (AMPS), developed by AT&T and Motorola, was the first

commercial cellular service in the United States and has been available to the public since 1983. This is known as a 1G analog system. In Europe, several similar 1G cellular systems have been deployed, including the total access communications system (TACS), the Nordic Mobile Telephone (NMT), C-450, the radio telephone mobile system (RTMS), and Radiocom. The Japanese TACS/narrowband TACS (JTACS/NTACS) are based on the European TACS system.

Analog wireless systems typically employ frequency modulation (FM) to modulate a carrier signal with voice information. The constant envelope feature of the FM signal enables high-efficiency amplification of the signal prior to transmission, since the power amplifier can be operated in the “saturated mode” without corrupting the information in the signal transmission. This advantage is historically exploited to lower the dc power, reduce the transistor cost, and reduce heat-sink requirements. Several analog cellular systems are listed in Table 1.1. These systems have FM and frequency-division multiple access (FDMA) in common.

In FDMA, multiple users are assigned different frequency channels, partitioned from the available frequency band. The use of FM and FDMA are the primary limitations on system capacity and user features for analog-based systems and have led cellular equipment manufacturers to adopt digital modulation techniques and alternative access methods. These are known as 2G systems. Digital modulation offers increased channel capacity, improved

Table 1.1
Several 1G Analog Wireless Systems

Standard	AMPS/NAMPS	ETACS	JTACS/NTACS
Year introduced	1983/1988	1985	1988/1993
Uplink frequency band (MHz)	North America 824–849	United Kingdom 890–915	Japan 915–925
Channel bandwidth (kHz)	30/10	25	25/12.5
Multiple access	FDMA	FDMA	FDMA
Modulation	FM	FM	FM
Maximum transmit power (dBm)	27.8	N/A	N/A
PA voltage (V)	3.6–6.0	3.6–6.0	3.6–6.0
Typical PA quiescent current (mA)	30	30	30
Typical efficiency (%)	>50	>50	>50

From: [6].

transmission quality, secure communication, digital data communication, and the ability to provide other value-added services not possible with analog modulation and FDMA.

Time-division multiple access (TDMA) is an access method in which multiple users share a common frequency band, with each user assigned a particular time slot. The most popular TDMA-based digital cellular system is the global system for mobile communications (GSM), which went into commercial service in 1992 [7]. GSM, which was originally developed as a Pan-European unified cellular standard to supplant the existing incompatible analog systems, has now become the most widely accepted wireless standard around the world. Gaussian minimum-frequency shift keying (GMSK) modulation has been adopted in GSM, which combines continuous-phase modulation with Gaussian-shaped filtering. GMSK is indeed a kind of frequency shift keying (FSK), which naturally results in phase changes of the carrier, while the carrier's magnitude remains constant. This technique, like FM, allows the power amplifier to operate in constant amplitude, or "saturation mode," and provides high-efficiency amplification.

In the United States, the Telecommunications Industry Association (TIA) adopted the IS-54 TDMA standard as an upgrade path for the analog AMPS to meet the growing need for increased cellular capacity in crowded areas. This standard later evolved with some improved services and support for personal communication services (PCS), and it is now referred to as IS-136. The other remaining TDMA wireless systems are the personal digital cellular (PDC) system and a microcellular standard—the personal handy-phone system (PHS) in Japan. These systems use a linear modulation technique, $\pi/4$ -differential quadrature phase shift keying (PSK) ($\pi/4$ -DQPSK) modulation, to provide better bandwidth efficiency and enable non-coherent detection. $\pi/4$ -DQPSK belongs to the PSK family of modulation, and the signal phase is intrinsically discontinuous, which causes the spectrum utilized to grow dramatically. Square-root raised cosine pulse shaping is used for band-limiting the resulting signal. The band-limited signal now has a significant variation in its amplitude, or envelope, which leads to the requirement for a linear power amplifier to accommodate the instantaneous envelope variation with minimal distortion.

For those systems that employ envelope-varying modulation schemes, a direct trade-off exists between the linearity and efficiency of the power amplifier. A power amplifier designed for these systems may also require dual-mode operation to maintain compatibility with any analog system that may coexist with the digital system.

While the radio spectrum is shared in the frequency domain in FDMA and shared in the time domain in TDMA, the code-division multiple access (CDMA) approach allows multiple users to occupy the same frequency spectrum at the same time. The separation of the users is based on code orthogonality; in other words, the separation is based in the code domain. Each user is assigned a unique identifying code, and, as a result, all transmitters except the desired user appear as additive white Gaussian noise at the intended receiver.

CDMA provides a number of advantages over its FDMA and TDMA counterparts, such as universal frequency reuse, increased capacity, use of a Rake receiver, different types of handoff, and accurate power control [8]. The commercial implementation of CDMA was carried out by Qualcomm, with the first cellular system operation in Hong Kong in 1995. This CDMA system, called IS-95 CDMA, uses offset quadrature PSK (OQPSK) for the mobile station transmitter. OQPSK is a variant of QPSK, which operates by delaying the signal by half a symbol period in the quadrature channel to avoid envelope zero-crossing and lessen the dynamic range requirements of the power amplifier. The band-limiting for the mobile station requires that the power amplifier for IS-95 CDMA be highly linear. Correspondingly, the power amplifier typically will exhibit low dc-to-RF conversion efficiency to support the large instantaneous envelope variation of the carrier. Furthermore, CDMA handset transmitters typically require a wide variation in the output power, due to the requirement for power control to combat the “near-far” problem in the network. This “near-far” phenomenon is due to the dispersion of multiple mobile users in the cell that share the same base station and spectrum. While the base station may receive differing signal power from each individual user, the users that deliver higher power unavoidably create excessive noise to others, and the system capacity is degraded. The cure to this problem is accurate power control, which equates the signal power delivered to the base station from each user. Table 1.2 lists several power amplifier-related features of 2G wireless standards.

1.2 Characterization of Power Amplifiers for Wireless Communications

In the hierarchy of the modern wireless communications transmitter system, the microwave linear power amplifier is the final interface between the baseband signal-processing RF upconversion and the antenna itself.

Table 1.2
Several 2G Digital Wireless Systems

Standard	GSM	IS-54	IS-95	PDC	PHS
Year introduced	1990	1991	1993	1991	1993
Uplink frequency band (MHz)	Europe 890–915	North America 824–849	North America 824–849	Japan 940–956	Japan 1,895–1,907
Carrier spacing (kHz)	200	30	1,250	25	300
Multiple access	TDMA/FDMA	TDMA/FDMA	CDMA/FDMA	TDMA/FDMA	TDMA/FDMA
Modulation	GMSK	$\pi/4$ -DQPSK	OQPSK	$\pi/4$ -DQPSK	$\pi/4$ -DQPSK
Duplex mode	FDD	FDD	FDD	FDD	TDD
Maximum transmit power (dBm)	30	27.8	27.8	33.0	19.0
Long-term mean power (dBm)	21.0	23.0	17.0	28.0	10.0
Peak-to-average power ratio (dB)	0	3.2	5.1	2.6	2.6
Transmit duty ratio (%)	12.5	33.3	Variable	33.3	33.3
PA voltage (V)	3.5–6.0	3.5–6.0	3.5–6.0	3.5–4.8	3.1–3.6
ACPR (dBc)	N/A	–26	–26	–48	–50
Typical PA quiescent current (mA)	20	180	200	150	100
Typical efficiency (%)	>50	>40	>30	>50	>50

From: [6]. FDD: frequency division duplexing; PA: power amplifier.

When viewed in this light, the power amplifier's function appears to be rather prosaic: a simple amplification of the input signal and delivery of the resulting power to the antenna. This is shown schematically in Figure 1.1 for the case of a typical wireless handset transmitter. However, this apparent simplicity masks the fact that the power amplifier often dominates the power dissipation in the handset and is the final determiner of the quality of the transmitted waveform. As a result, a careful analysis of this simple block reveals a host of interesting problems. Most of the problems are associated with the quality of the transmitted waveform, and the dc power dissipation of the amplifier itself.

1.2.1 Power Amplifier Waveform Quality Measurements

Table 1.2 lists the relevant power amplifier characteristics for several 2G digital wireless standards. An important distinction between most digital standards and the analog standards is the inclusion of a transmitter linearity specification, which specifies the leakage or interference (or both) to the neighboring channels. This interference specification is required, because the nonlinear distortion of the amplifier creates an output spectrum that is broader than the original input signal. This broadening of the transmitted signal can interfere with users in nearby channels. How is this distortion created, and how is it affected by the characteristics of the power amplifier and the characteristics of the signal sent through the power amplifier?

The nonlinear distortion of the amplifier usually results from nonlinear distortion processes in the transistors that make up the amplifier. Many

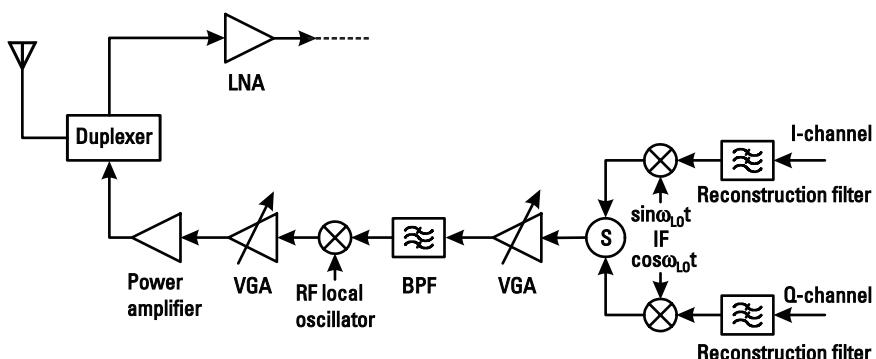


Figure 1.1 Typical wireless handset transmitter section, showing complete upconversion path from the baseband.

high-frequency circuits like power amplifiers can be roughly characterized by assuming that they exhibit *memoryless weak* nonlinearities. In this case, they can be accurately characterized through the use of a *power series* expansion. Therefore, the instantaneous output of a circuit can be represented by

$$s_o(t) = a_1 s_i(t) + a_2 s_i^2(t) + a_3 s_i^3(t) + \dots \quad (1.1)$$

where $s_o(t)$ is the output of the circuit, $s_i(t)$ is the input to the circuit, and $a_1, a_2, a_3 \dots$ are the power series coefficients of the output. The quantity a_1 corresponds to the linear gain coefficient of the circuit, and $a_2 \dots a_n$ represent the nonlinearities of the circuit due to nonideal elements such as power supply limitations. Figure 1.2 is a general illustration of the nonlinear behavior of the circuit.

We can immediately see from this expression several different consequences of the nonlinear behavior of the circuit. A single sinusoidal input signal at frequency ω_1 will generate outputs at frequencies $\omega_1, 2\omega_1, 3\omega_1, \dots n\omega_1$ for an n th order nonlinearity. Of course, a typical transmitted waveform contains a multitude of frequencies, and the frequency generation that results from this condition is even more complicated. In order to capture this behavior in a straightforward way, communications engineers have developed *two-tone* tests to model the behavior of the circuit. In this case, the input to the circuit consists of two sinusoidal signals, such as

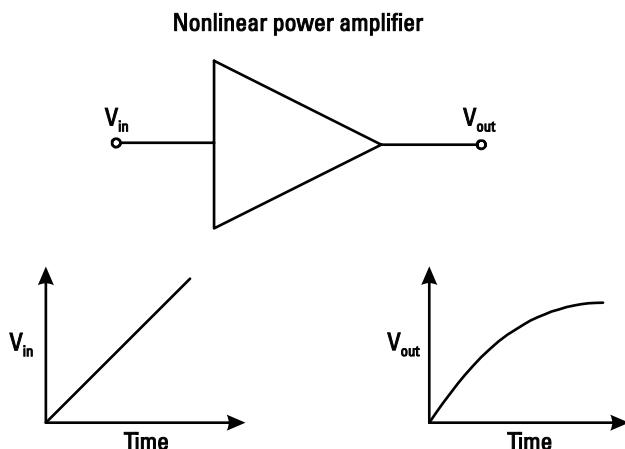


Figure 1.2 Illustration of general power amplifier nonlinearity characterized by a power series.

$$s_i(t) = S_1 \cos(\omega_1 t) + S_2 \cos(\omega_2 t) \quad (1.2)$$

Taking only the first three terms of the power series expansion ($a_4, a_5, \dots = 0$), and substituting in the appropriate trigonometric identity will result in the output

$$s_o(t) = a_1 s_i(t) + a_2 s_i^2(t) + a_3 s_i^3(t) \quad (1.3)$$

$$\begin{aligned} s_o(t) = & \overbrace{(1/2)a_2(S_1^2 + S_2^2)}^{\text{dc term}} \\ & + \overbrace{[a_1 S_1 + (3/4)a_3(S_1^3 + 2S_1 S_2^2)] \cos(\omega_1 t)}^{\text{output at desired frequency}} \\ & + \overbrace{[a_1 S_2 + (3/4)a_3(S_2^3 + 2S_2 S_1^2)] \cos(\omega_2 t)}^{\text{output at desired frequency}} \\ & + \overbrace{a_2 S_1 S_2 [\cos(\omega_1 + \omega_2)t + \cos(\omega_1 - \omega_2)t]}^{\text{intermodulation term}} \\ & + \overbrace{(3/4)a_3 S_1^2 S_2 \cos(2\omega_1 - \omega_2)t}^{\text{intermodulation term}} \\ & + \overbrace{(3/4)a_3 S_2^2 S_1 \cos(2\omega_2 - \omega_1)t}^{\text{intermodulation term}} \\ & + \overbrace{(1/2)a_2[S_1^2 \cos(2\omega_1 t) + S_2^2 \cos(2\omega_2 t)]}^{2 \times \text{frequency term}} \\ & + \overbrace{(3/4)a_3 S_1^2 S_2 \cos(2\omega_1 + \omega_2)t}^{\text{intermodulation term}} \\ & + \overbrace{(3/4)a_3 S_2^2 S_1 \cos(2\omega_2 + \omega_1)t}^{\text{intermodulation term}} \\ & + \overbrace{(1/4)a_3[S_1^3 \cos(3\omega_1 t) + S_2^3 \cos(3\omega_2 t)]}^{3 \times \text{frequency term}} \end{aligned} \quad (1.4)$$

Further expansion of the power series at higher orders is a straightforward, if frustrating, procedure; it is best left to a computer in most cases. A simplified plot of the magnitude and frequencies of these

output terms appears in Figure 1.3. Fortunately, higher terms of the power series are usually of little consequence in a practical circuit, and a variety of very useful results can be seen from this simplified two-tone third-order result.

Equation (1.4) demonstrates a variety of interesting consequences of the nonlinear distortion of the input signal. They are listed as follows:

- Change in the bias point of the circuit;
- Gain compression, or expansion, at the desired frequency depending on the sign of a_3 ;
- Creation of harmonics of the original input frequencies at $2\omega_1$, $3\omega_1$, $2\omega_2$, and $3\omega_2$;
- Creation of frequencies that are linear combinations of the original input frequencies. These *intermodulation* frequencies are at $2\omega_1 \pm \omega_2$, $2\omega_2 \pm \omega_1$, and $\omega_2 \pm \omega_1$.

Note that the output frequencies generated by the nonlinearity are a linear combination of the input frequencies. In general, this means that

$$\omega_o = a\omega_1 + b\omega_2 + c\omega_3 + \dots d\omega_k \quad (1.5)$$

where a , b , c , and d are integers from $-n \dots n$, and n is the maximum order of the nonlinearity— $n = 3$ in this case. The input frequencies to the circuit are $\omega_1, \omega_2 \dots \omega_k$, and the output frequency is ω_o .

The change in the bias point of the circuit is usually small and is a consequence of the fact that the second-order (and all additional even-order)

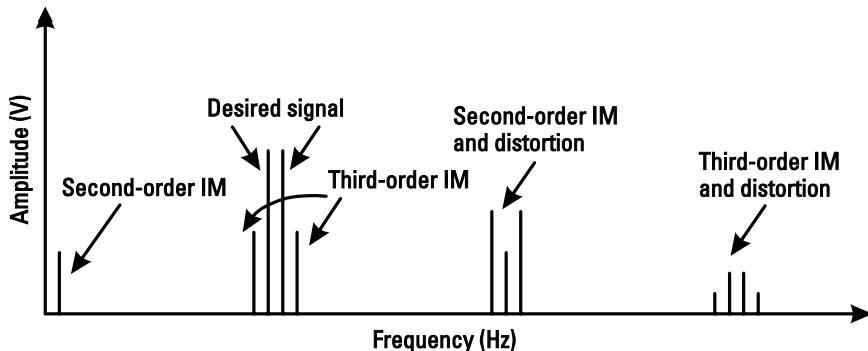


Figure 1.3 Illustration of harmonic and intermodulation distortion products in two-tone amplifier measurements.

harmonic distortion generators create a *dc* term in the output. The compression or expansion of the gain of the linear gain of the circuit is attributed to the a_3 term in (1.1), which has a negative value for a typical compressive nonlinearity. This effect is usually characterized by the 1-dB compression point of the power amplifier—the signal level where the gain of the circuit is reduced by 1 dB.

The magnitude of the 1-dB compression point can be estimated from (1.4). If we assume that S_2 is zero then the output signal at frequency ω_1 will be

$$s_o(t) = [a_1 S_1 + (3/4)a_3 S_1^3] \cos(\omega_1 t) \quad (1.6)$$

and the relative gain compared to the desired (uncompressed) gain, at the -1-dB compression point, is

$$-1 = 20 \log \frac{a_1 S_1 + (3/4)a_3 S_1^3}{a_1 S_1} \quad (1.7)$$

and solving for S_1 yields a value for the input -1-dB compression point of

$$S_{-1\text{dB}} = \left[\frac{4}{3} \frac{a_1}{a_3} \right]^{1/2} \sqrt{0.11} \quad (1.8)$$

This expression can be used to calculate the 1-dB compression point of the amplifier if the power series coefficients for the gain are known.

A related aspect of the performance of the power amplifier is AM/AM and AM/phase modulation (PM) conversion. AM/AM conversion is the change in the gain of the circuit with changes in the input amplitude. AM/PM conversion is an aspect of nonlinear circuit behavior where changes in the *amplitude* of the input signal as it is applied to a circuit cause a change in the *phase* of the signal at the output. This can be especially detrimental to phase-modulated digital waveforms, where the bit error rate (BER) can rise as a result of changes in the phase of the received or transmitted signal. This phenomenon can not be predicted by a memoryless nonlinearity power-series of (1.1), which intrinsically contains no phase information.

We can gain a simple understanding of the effect by assuming a single-tone input signal and by assuming that a_1, a_2, \dots, a_n are *phasor* rather than *scalar* quantities, each consisting of magnitude $|a_n|$ and phase θ_n . This turns out to be a good approximation to the results that would be obtained from an exact Volterra-series analysis. Now, the output vector at the desired frequency is [from (1.6)]

$$s_o(t) = \left(|a_1| e^{j\theta_1} S_1 + (3/4) |a_3| e^{j\theta_3} S_1^3 \right) \cos(\omega_1 t) \quad (1.9)$$

Clearly, the overall output phase of this result will vary along with the output amplitude as the amplitude of the input signal S_1 increases, resulting in AM/PM conversion. Thus, AM/PM conversion is a consequence of the phase relationships between the power series coefficients.

Although they both arise from the same phenomenon, engineers have developed a distinction between *harmonic* distortion, which generates output frequencies at integral multiples of the input frequencies, and *intermodulation* distortion, which generates output frequencies at linear combinations of integral multiples of the input frequencies. In this way, two input signals are said to *intermodulate*. It is this second set of distortion products that is particularly insidious in communications systems, since the distortion product is often in the same frequency band as the desired input signals and is therefore indistinguishable from an actual signal. It is clear from (1.4) that the magnitude of the harmonic and intermodulation distortion terms are related.

Intermodulation distortion is typically characterized with a two-tone test, at conditions where $S_1 = S_2$. As a result, the *second-order* intermodulation products are

$$s_o(t) = a_2 S_1^2 [\cos(\omega_1 + \omega_2)t + \cos(\omega_1 - \omega_2)t] \quad (1.10)$$

and the *third-order* intermodulation products are

$$s_o(t) = (3/4) a_3 S_1^3 [\cos(2\omega_1 - \omega_2)t + \cos(2\omega_2 - \omega_1)t] \quad (1.11)$$

and we can define the fractional second-order intermodulation (IM_2) as

$$\begin{aligned} \text{IM}_2 &= \frac{\text{amplitude of the second-order intermodulation output}}{\text{amplitude of the fundamental output}} \\ &= \frac{|a_2| S_1^2}{|a_1| S_1} \\ &= \frac{|a_2| S_1}{|a_1|} \end{aligned} \quad (1.12)$$

and the fractional third-order intermodulation (IM_3) is

$$\begin{aligned}
 \text{IM}_3 &= \frac{\text{amplitude of the third-order intermodulation output}}{\text{amplitude of the fundamental output}} \\
 &= \frac{3|a_3|S_1^3}{4|a_1|S_1} \\
 &= \frac{3|a_3|S_1^2}{4|a_1|} \tag{1.13}
 \end{aligned}$$

Notice that the value of IM_2 rises linearly with the input signal, so a 1-dB increase in S_1 results in a 1-dB increase in IM_2 . Similarly, the value of IM_3 rises linearly as the square of the input signal, so a 1-dB increase in S_1 results in a 2-dB increase in IM_3 .

Now, in the performance of a two-tone test, with only second- and third-order nonlinearities present, the desired output signal is at frequencies ω_1 and ω_2 , and the undesired output frequencies occur at $\omega_1 + \omega_2$ and $\omega_1 - \omega_2$ (for second-order intermodulation) and $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$ (for third-order intermodulation). We can plot the relative value of the desired output, the second-order intermodulation output, and the third-order intermodulation output for any given nonlinearity as a function of input signal amplitude, as shown in Figure 1.4. Note that these quantities are typically plotted on a log-log scale to make the quantities linear functions of the input. In this case, the desired signal amplitude has a slope of one; the second-order intermodulation product has a slope of two; and the third-order

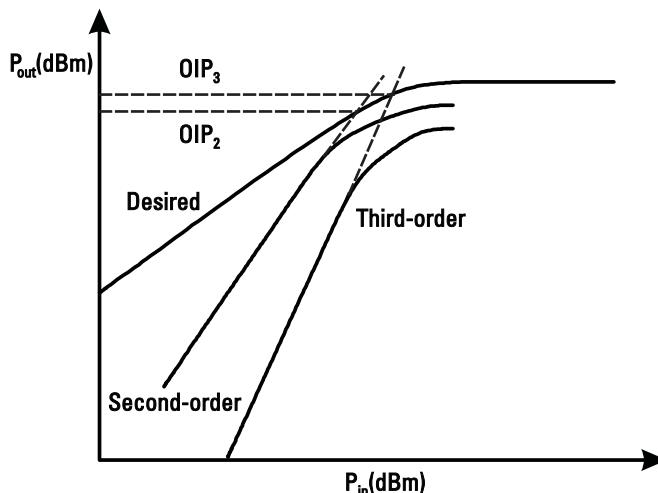


Figure 1.4 Illustration of extrapolated nonlinear amplifier intercept points.

product has a slope of three. Since the slopes of the intermodulation products are of necessity higher than that of the desired signal level, there will be some extrapolated point where they are *equal* to the input signal level. This is shown in Figure 1.4.

The input signal level where the extrapolated value of the desired output signal and the n th-order intermodulation product are equal is the n th order input intercept point- (S_{IIP_n}) . So by definition S_{IIP_n} is defined as the input signal amplitude where $IM_n = 1$. So, from (1.12) and (1.13).

$$S_{IIP_2} = \frac{|\alpha_1|}{|\alpha_2|} \quad (1.14)$$

and

$$S_{IIP_3} = \sqrt{\frac{4|\alpha_1|}{3|\alpha_3|}} \quad (1.15)$$

This last expression can be used to quantify the well-known relationship between the 1-dB compression point (1.6) and the third-order intercept point; for example,

$$S_{IIP_3}(\text{dB}) = S_{-1\text{dB}}(\text{dB}) + 9.6 \quad (1.16)$$

In high-power amplifier applications, it is often more useful to specify system performance in terms of *output* intercept point (OIP) rather than IIP. These can be calculated in a straightforward manner from (1.15) and (1.14) as

$$S_{OIP_2} = \frac{|\alpha_1|^2}{|\alpha_2|} \quad (1.17)$$

and

$$S_{OIP_3} = \sqrt{\frac{4|\alpha_1|^3}{3|\alpha_3|}} \quad (1.18)$$

Now, the two-tone test approach combined with a weak nonlinearity is a useful starting point for the understanding of nonlinear power amplifier behavior. However, typical power amplifiers in digital wireless applications are characterized by *strong* nonlinearities and *digital modulation*. Under these constraints, exact expressions for spectral regrowth are far more complex.

These conditions lead to a slightly different, though closely related, set of performance constraints as shown in the following simple example.

A large-signal nonlinear power amplifier can be characterized most simply by a clipping nonlinearity, as shown in Figure 1.5. In this case, the amplifier response is perfectly linear, with a gain of unity, until the saturation point is reached, beyond which the output grows no further.

Now, the effect of this clipping nonlinearity on the resulting modulated output spectrum is quite dramatic, as shown in Figure 1.6. In this case, an 8-PSK waveform, with a raised cosine filter ($\alpha = 0.35$) and an approximately 1 Ms/s symbol rate is sent to the amplifier input with an available power of 0 dBm. As Figure 1.6 demonstrates, the narrow bandwidth of the input waveform is effectively destroyed by the intermodulation resulting from the clipping nonlinearity of the amplifier.

This effect can also be seen by an examination of the signal constellation diagram of the input and output waveforms (Figure 1.7). In Figure 1.7(a), the ideal input signal constellation exhibits wild variation in its amplitude in the time domain, but the resulting spectrum is quite narrow in the frequency domain. The limiting of the output amplitude in Figure 1.7(b) results in a dramatic constriction of the signal in the time domain and a corresponding spectral regrowth.

The signal that is sent through the amplifier can be characterized by the peak-to-average power (PAP) ratio and the complementary cumulative distribution function (CCDF), which are statistical measurements on the envelope of the transmitted time-domain waveform. The PAP ratio is the ratio of the peak envelope power of the waveform to the average envelope power during a set period of time (usually as long as possible). In some cases,

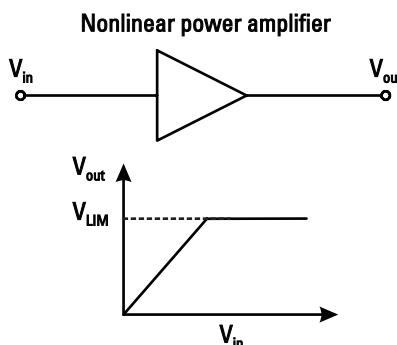


Figure 1.5 Simple power amplifier exhibiting clipping nonlinearity. In this case, the output voltage saturates at V_{LIM} .

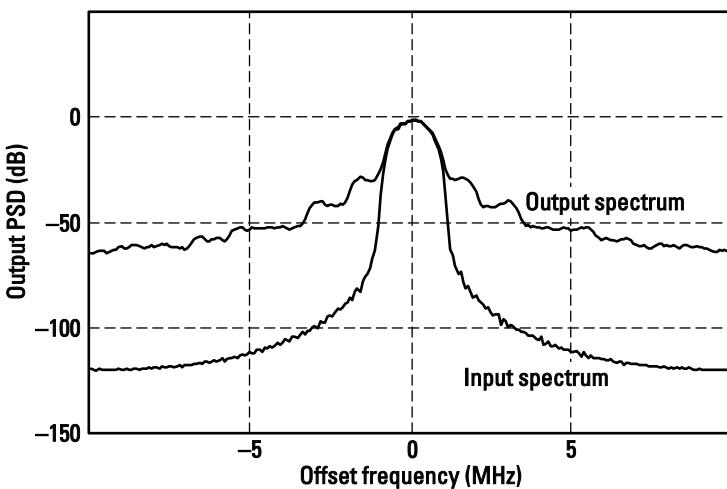
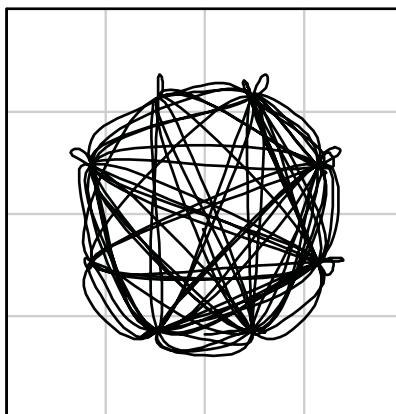


Figure 1.6 Input and output spectrum of limiting power amplifier. Note the out-of-band spectral regrowth resulting from the limiting behavior.

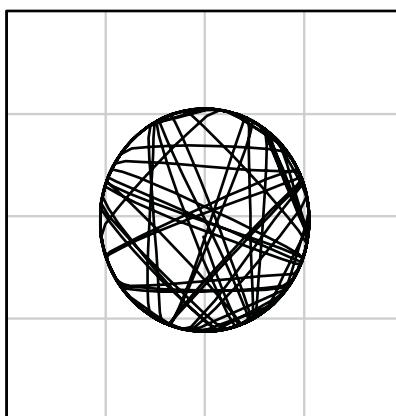
it is not possible to determine the peak of the waveform definitively, since the signal has a very wide distribution of possible amplitudes, and then a probabilistic measurement of the PAP is used. For example, the peak envelope power in this case is not specified as an *absolute* peak, but rather as the power level that the signal is *below* for a certain percentage of the time—typically 99.9 or 99.99% of the time.

The power statistics of the signal can also be characterized graphically by the CCDF. The CCDF curve shows the probability that the power is equal to or above a certain PAP ratio. The PAP ratio and CCDF plots are useful characterization techniques in digital communication transmitters, since the modulation formats vary widely. As an example, in IS-95 CDMA systems, the statistics of the signal will be dependent on how many code channels /or carriers (or both) are present at the same time. Figure 1.8 shows the CCDF curves with different code-channel configurations [9]. Even in systems that use constant-amplitude modulation—such as GSM—the PAP ratio can be greater than unity if the transmitter is amplifying more than one signal, such as in base stations.

Given the fact that the modulated signal, with a given CCDF and PAP, is passed through the nonlinear power amplifier, the linearity figure of merit for digital wireless communication systems is typically the ACPR and the alternate channel power regrowth (AltCPR). The ACPR is typically defined as the ratio of the distortion power measured within a specified bandwidth



In-phase component
(a)



In-phase component
(b)

Figure 1.7 Signal constellation diagrams for (a) ideal filtered 8-PSK waveform, and (b) "clipped" waveform illustrating constriction of constellation diagram after passing through a limiting amplifier.

B_{out} in the adjacent channel at a specified offset frequency from the channel center frequency f_c , to the signal power measured around the center frequency within another specified bandwidth B_{in} in the desired channel; the Alt CPR is a measure of the ratio of distortion power in the alternate channel to the signal power in the desired channel. These two measures are shown in Figure 1.9. The two bandwidths B_{out} and B_{in} are different in many cases. The spectral regrowth results from the power amplifier nonlinearities

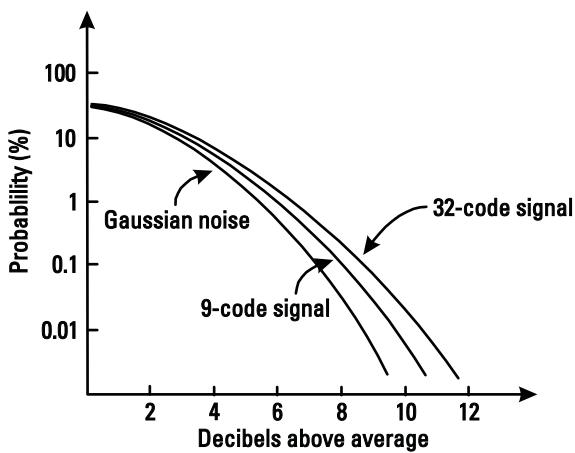


Figure 1.8 CCDF of a CDMA signal with differing codes and a comparison to a Gaussian noise profile [9].

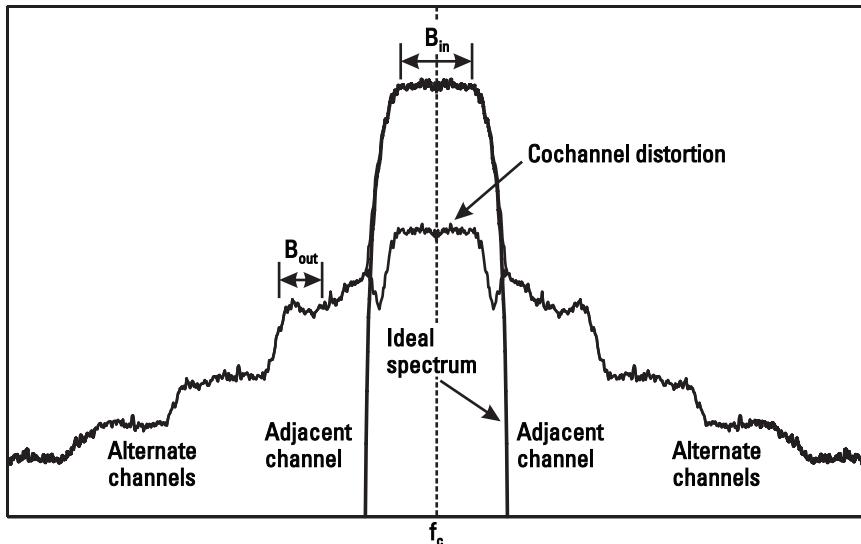


Figure 1.9 Spectrum of the ideal transmitted modulated signal and the distorted signal illustrating ACPR, alternate channel regrowth, and inband distortion.

(AM-AM and AM-PM conversion) and will cause interference to the users in the adjacent and alternate channels. In general, the third-order nonlinearities in the amplifier contribute to the adjacent channel spectral regrowth, and

the fifth-order nonlinearities create both the adjacent and alternate channel spectral leakage. Figure 1.9 also illustrates inband signal distortion.

Each of the digital standards characterizes ACPR requirements differently; they are usually specified by an RF spectrum mask that is related to the spacing between channels, as well as limitations on the out-of-band emissions specified by the regulating body. In addition, each standard has its own unique modulation format, and the distribution function of the waveform can alter the ACPR. While the maximum adjacent and alternate channel powers are usually specified by wireless standards, in practice a more straightforward and convenient measure for the purpose of simulation of amplifier linearity is the adjacent channel interference (ACI). In this text, the ACI is defined as the ratio of the peak spectral density of the residue outside the channel to the peak spectral density of the modulation.

There are alternative techniques for characterizing power amplifier accuracy and linearity in modern communications systems. The previous figures of merit characterized the *spectral regrowth* of the signal. Other figures of merit characterize the *accuracy* of the modulated signal. They typically involve a precision demodulation of the transmitted signal and subsequent comparison to an ideal reference signal. The figure of merit depends mainly on the modulation scheme and the wireless standard. The NADC and PDC systems use the error vector magnitude (EVM) measurement, while GSM uses phase and frequency error. The CDMA IS95 system employs a waveform quality metric ρ .

The EVM measurement is a modulation quality metric widely used in digital RF communications systems, especially emerging third generation (3G) and wireless local area networks. It is essentially a measure of the accuracy of the modulation of the *transmitted* waveform. Mathematically, the EVM is defined as [10]

$$\text{EVM} = \frac{\sqrt{\sum_n |e(k)|^2}}{n} \quad (1.19)$$

where $e(k)$ is the normalized magnitude of the error vector at symbol time k , and n is the number of samples over which the measurement is made. Typical EVM figures are in the range of 5–15%.

Another way of looking at this is that the EVM is the root-mean-square (rms) value of the error vector when the symbol clock transitions occur. This is shown in simplified form in Figure 1.10. The error vector is a complex quantity, containing both magnitude and phase components. For this reason,

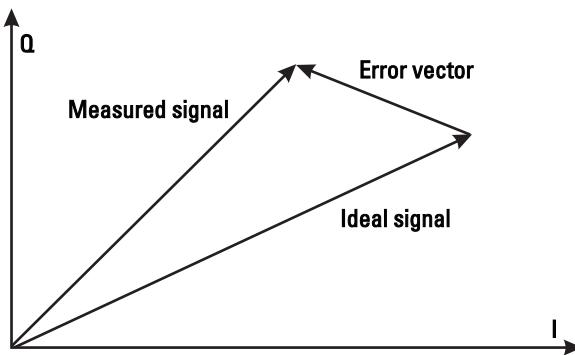


Figure 1.10 EVM is a measure of the difference between the ideal and actual transmitted waveforms.

when the power amplifier input signal has a small amount of distortion and noise, there is a simple relationship between the EVM of the amplifier output and the AM-AM and AM-PM characteristic [11].

The waveform quality metric (ρ) is yet another measure of the fidelity of the transmitted signal, which is typically used in CDMA systems. In this case, it is the cross correlation of the transmitted to the ideal baseband signal, as follows

$$\rho = \frac{\sum_{k=1}^M D_k S_k}{\sqrt{\sum_{k=1}^M D_k \cdot \sum_{k=1}^M S_k}} \quad (1.20)$$

where S_k is the k th sample of the transmitted signal, D_k is the k th sample of the ideal baseband signal, and M is the measurement period in half-chip intervals. In most cases, the waveform quality factor usually measures about or above 0.98 or better.

1.2.2 Power Efficiency Measurements

A comparison of the quiescent current values of Table 1.1 and Table 1.2 shows that linear power amplifiers require a substantially higher quiescent current than that used for constant envelope applications. This characteristic, coupled with the requirement that linear power amplifiers cannot be driven into deep saturation, are the two reasons why linear power amplifiers exhibit lower efficiency than those for constant envelope applications. The 3G

wireless standards like wideband CDMA (WCDMA) and CDMA2000 demand even more strict requirements on the power amplifier linearity performance [12].

One related problem is that the dc-to-RF amplifier efficiency generally drops sharply as the RF input drive power is “backed-off” from the maximum rated power level. The power-added efficiency (PAE) of a typical amplifier is a measure of the conversion efficiency of all sources of input power (both from the power supply and the input signal) to the output, and is given by

$$\text{PAE} = \eta = \frac{P_{\text{out}}}{P_{\text{in}} + P_{\text{dc}}} \quad (1.21)$$

where P_{in} is the RF input power to the amplifier, P_{out} is the desired output power of the amplifier in the band of interest, and P_{dc} is the dc input power supplied to the circuit. If the gain of the circuit is relatively high, then the RF input power is much smaller than the dc power, and the PAE is a measure of the conversion efficiency from the battery to the transmitted signal.

Now, the output power delivered by the amplifier varies as a function of numerous factors, including the position of the mobile unit within the network. In CDMA networks, this variation is due to the fact that the received signal strength must be maintained at a *constant* level at the base station to combat the near-far problem. The dc power drawn by the amplifier naturally varies as the RF output power varies, and so a more useful figure of merit for these amplifiers is the long-term mean efficiency given by

$$\bar{\eta} = \frac{\int_{-\infty}^{\infty} P_{\text{out}} \cdot g(P_{\text{out}}) dP_{\text{out}}}{\int_{-\infty}^{\infty} P_{\text{dc}}(P_{\text{out}}) \cdot g(P_{\text{out}}) dP_{\text{out}}} \quad (1.22)$$

where $g(P_{\text{out}})$ is the probability that the amplifier will have an output power P_{out} , and $P_{\text{dc}}(P_{\text{out}})$ is the dc power dissipation at output power P_{out} .

In light of its inherent simplicity, it is not clear from (1.22) that the long term PAE of a typical linear power amplifier could not approach 100% in a realistic situation. This would represent an ideal situation where nearly all of the dc power were converted to transmitted power. Unfortunately however, typical power amplifiers in linear transmitter applications have average PAEs averaging 5% or less [13], and the numbers are equally poor for base station applications.

The origins of this poor efficiency in practical situations are straightforward. Consider the case of a simple bipolar transistor power amplifier shown in Figure 1.11. In this case, the current-voltage characteristics dictate that the dc collector bias on the device is V_{cc} and the dc current through the device is $I_{max}/2$. As a result, the dc power dissipation of the amplifier is simply $V_{cc}I_{max}/2$. Under conditions of maximum output power, the collector voltage swings from 0 to $2V_{cc}$ and the device current swings from 0 to I_{max} and delivers the maximum output power to the load impedance of $V_{cc}I_{max}/4$. Thus, the maximum efficiency of the amplifier under these conditions is 50%, and the load impedance that is presented to the device is $2V_{cc}/I_{max}$. In the example, the load impedance presented to the device is 5Ω .

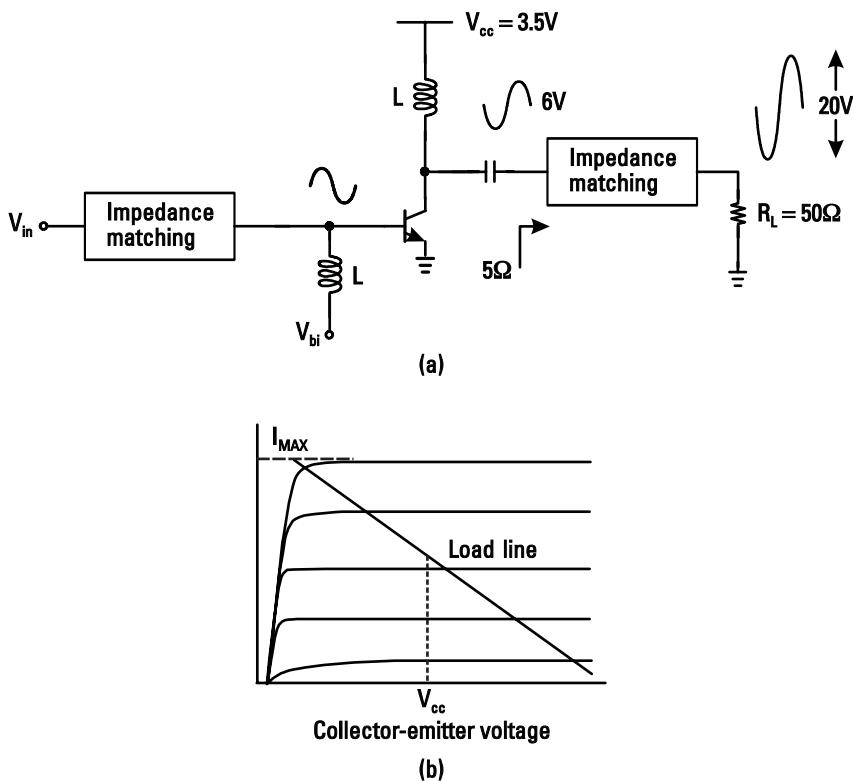


Figure 1.11 Simplified view of a power amplifier illustrating (a) circuit schematic with impedance matching at input and output to maximize the output power, and (b) typical transistor characteristics and load line.

However, in most cases, the actual efficiency is substantially less than this maximum amount, and a good portion of the dc power delivered to the power amplifier remains inside it, becoming waste heat. The lower efficiency results from the fact that the dc power dissipated by the amplifier essentially remains at $V_{cc}I_{max}/2$ for *all* values of output power, so the efficiency decreases linearly with output power.

1.3 Power Amplifier Linearization and Efficiency-Enhancement Techniques

What follows next is a brief overview of linearization and efficiency-enhancement techniques for microwave power amplifiers, followed by a historical overview of the outphasing power amplifier technique. The outphasing approach has a number of distinct advantages compared with other approaches, although the widespread adoption of the technique has required the development of modern digital-signal processing techniques, which make some of the more sophisticated digital control algorithms a practical reality.

Over the years, many power amplifier linearization and efficiency-enhancement techniques have been proposed. Some implementations take both linearization and efficiency-enhancement techniques together, while some specific techniques can be implemented with either linearization *or* efficiency enhancement as the primary goal. The most frequently discussed linearization techniques include Cartesian feedback, simple predistortion, adaptive digital predistortion, feed-forward, and outphasing amplifier. The efficiency-enhancement techniques comprise the Doherty amplifier, envelope elimination and restoration (EER), and bias adaption. Those techniques are well documented in the literature and have important and useful applications in modern wireless communications systems, for both mobile terminals and base station power amplifiers.

One of the simplest approaches for the improvement of linearity in the transmitter power amplifier is the well-known technique of predistortion. A typical power amplifier exhibits gain compression at high input powers, which results in AM-AM distortion. It also exhibits excess phase shift at high input powers, which results in AM-PM conversion. Together, these effects create distortion and intermodulation in the high-power output of the amplifier, hurting the ACPR and EVM performance.

If the input signal to the power amplifier could be “predistorted” with the inverse of its own nonlinearity, then the overall effect of the nonlinearity

could be canceled out. This is shown conceptually in Figure 1.12, where an analog or RF predistortion circuit compensates for both the gain and phase nonlinearity of the amplifier circuit. The predistortion circuit would typically exhibit both gain and phase *expansion* at high-input power levels, since a typical power amplifier exhibits gain and phase *compression* at high-output power levels.

Although straightforward in principle, the predistortion approach suffers from several practical drawbacks. First, it is very difficult to precisely track the effects of temperature, process, and power supply variations on the characteristics of the power amplifier nonlinearity. This is a serious drawback, because, as the previous sections showed, the amount of acceptable distortion in a typical 2G or 3G system is very low, and a small offset in the characteristics of the power amplifier and the predistortion circuit can create substantial out-of-band interference.

Another possible approach is to perform the predistortion using digital techniques at baseband frequencies, if the appropriate transformation characteristic for the predistorter were known in advance. This technique is illustrated in Figure 1.13 and is known as *adaptive predistortion* [14]. In this case, the AM-AM and AM-PM distortion through the amplifier is “measured,” and this data is then fed to a digital signal processor that provides the appropriate predistorted in-phase and quadrature-phase signals for the baseband upconverter. Of course, the problem is that the ideal transfer characteristic for the predistorter varies with time, and so the algorithm performing the predistortion must be periodically updated. Several different versions of adaptive predistortion have been developed [15].

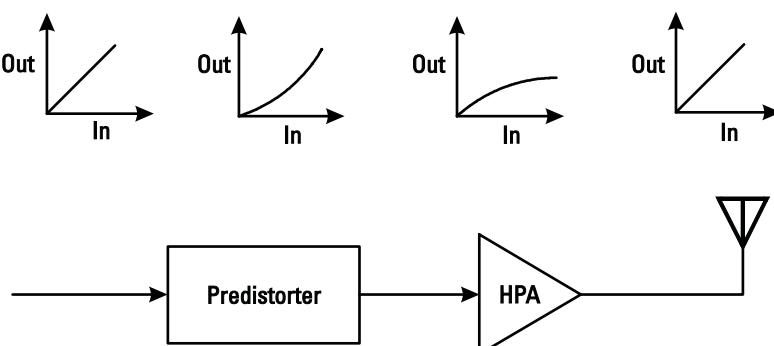


Figure 1.12 The predistortion concept works by adding a series inverse nonlinearity to the power amplifier. The combination of the two creates a linear input/output transfer characteristic.

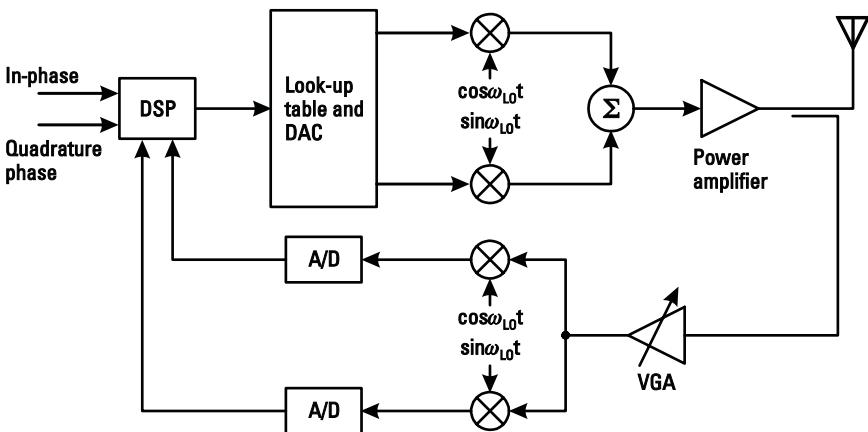


Figure 1.13 Adaptive predistortion employs a measurement of the output waveform to produce the necessary input compensation.

The practical limitations of the predistortion concept have naturally led to the development of more robust techniques for achieving power amplifier linearization. The traditional linearization technique for nonlinear analog systems is linear feedback. With appropriate feedback, the loop itself naturally compensates for the nonlinear transfer characteristic of the nonlinear power amplifier. An example of a hypothetical linear feedback approach for a power amplifier is illustrated in Figure 1.14(a). In this system, an operational amplifier supplies the necessary “predistortion” of the signal in response to the difference between the (distorted) output signal and the desired input signal. This straightforward approach has the obvious limitation that operational amplifiers with the required bandwidth and drive capability do not exist at microwave frequencies. Furthermore, the phase shift associated with a typical power amplifier is highly variable, making unconditional stability of the feedback circuit difficult to achieve under a wide range of conditions.

Providing the feedback at lower frequencies, where operational amplifiers have sufficient bandwidth, by downconverting the amplified signal to baseband frequencies is one possibility, as shown in Figure 1.14(b). The drawback of this approach is that the downconversion mixers have to be as linear as the desired output signal. This is not a problem in most cases, since only a small portion of the output signal is required for feedback purposes. Another problem is the excess and variable phase shift through the power amplifier, downconversion mixer, and lowpass filter combination, which is hard to control at microwave frequencies, and varies depending on

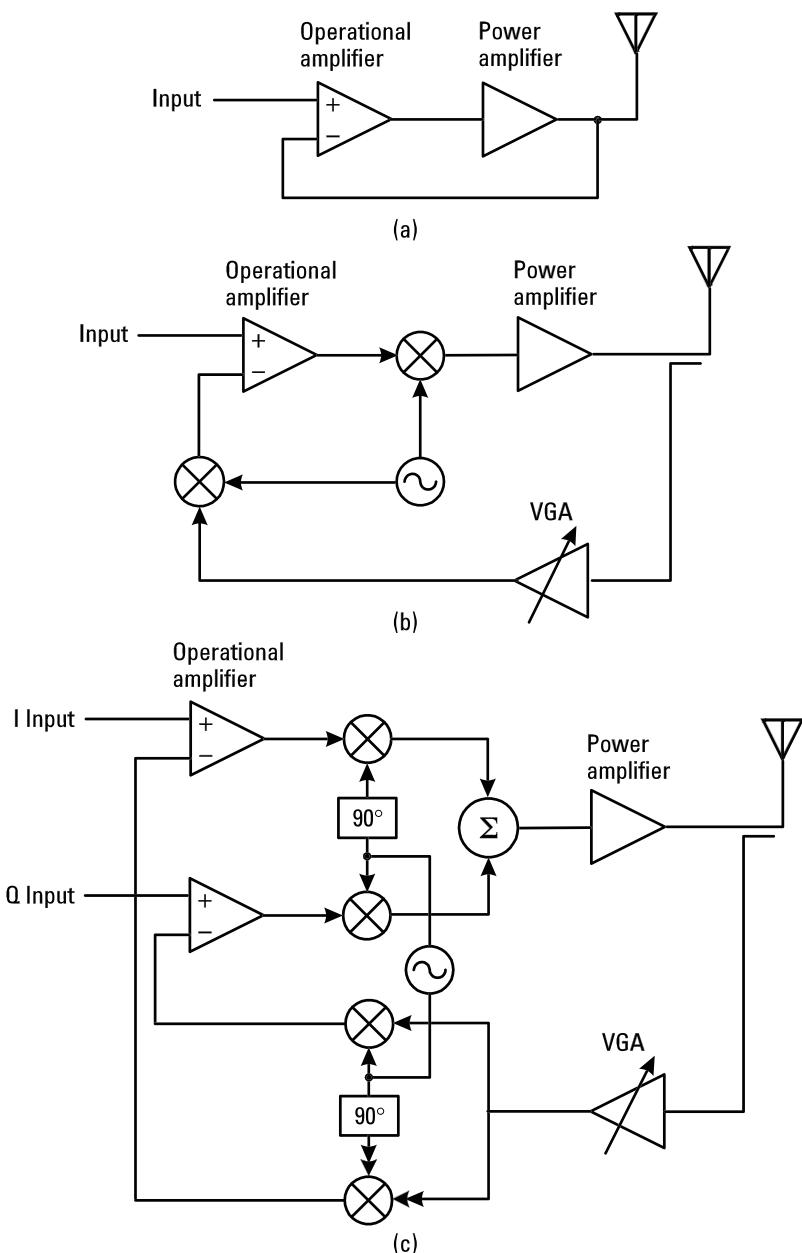


Figure 1.14 Amplifier linearization using feedback: (a) simplified view of feedback linearization approach, (b) use of frequency-translating downconverter to achieve linearization, and (c) Cartesian feedback applied to provide both gain and phase correction.

the power level. An additional phase shift is therefore often added to the input to the mixer to ensure stability under all conditions, and this phase shift must be carefully controlled over process, temperature, and power supply variations. The feedback approach is also prone to problems associated with amplifier saturation and rapid changes in output voltage standing-wave ratio (VSWR) [16]. Digital modulation techniques typically require upconversion of both the I and Q baseband signals. As a result, feedback is typically applied to both paths of the power amplifier inputs, with a technique known as Cartesian feedback, shown in Figure 1.14(c). Cartesian feedback has been an active research topic for many years [16], but has not yet achieved widespread adoption.

The myriad of problems associated with the predistortion approaches—both open-loop and feedback—point to an opportunity for alternative solutions. Rather than predistorting the input signal, it might be more effective to measure the nonlinearity of the power amplifier, subtract the error generated by the nonlinearity from the “ideal” signal, amplify the difference, and then subtract that difference (which is the error) from the amplifier output. The result would be an “error-free” amplification of the input signal. This approach, although seemingly complicated, has been used successfully for many years to linearize satellite traveling wave tube amplifiers (TWTAs) and is known as the feed-forward approach [17]. It is illustrated schematically in Figure 1.15.

Feedforward techniques for amplifier linearization actually predate the use of feedback techniques; both were developed by Black in the 1930s to solve the problem of linearization for telephone network repeater amplifiers [18]. A close examination of Figure 1.15 illustrates the reason that feedback techniques quickly supplanted those of feed-forward for most lower frequency applications. First of all, the gain and phase matching between the two input paths of the subtractor circuit must be very precisely matched to achieve acceptable cancellation of the distortion products. Second, the gain of the error amplifier must precisely track the gain of the power amplifier itself. Finally, the phase shift through the final phase shift network and hybrid coupler must precisely track the gain and phase shift of the power amplifier. Despite these apparent obstacles, the use of feed-forward approaches has several advocates, although it is typically employed in base station and higher frequency circuits, where power efficiency is less important than absolute linearity.

All of these various techniques—and their myriad limitations—highlight the need for yet another improved strategy for the linearization of power amplifiers. The outphasing approach is a technique whose origin

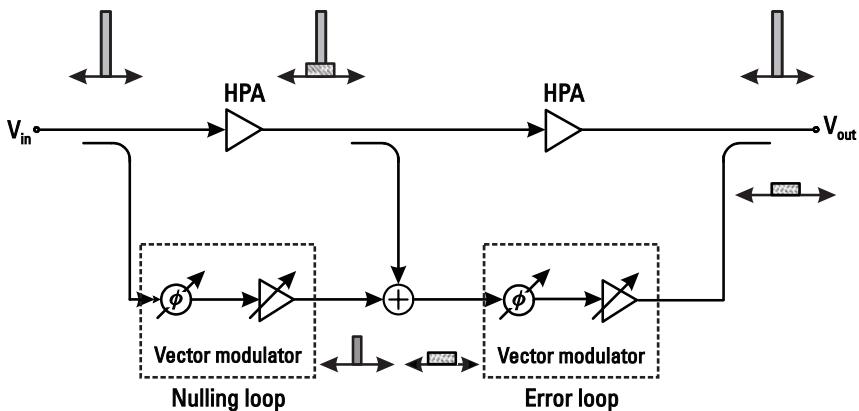


Figure 1.15 Feed-forward predistortion of nonlinear power amplifiers.

dates to the 1930s but that has undergone something of a renaissance in recent years with advances in DSP and integrated circuit technology.

1.4 Outphasing Microwave Power Amplifiers

The concept of outphasing amplification is a technique with a long history (dating to Chireix in the 1930s [19]) that has been revived under the rubric of linear amplification with nonlinear components (LINC) and applied to a variety of wireless applications [20–22]. A variation on the approach is known as combined analog locked-loop universal modulator (CALLUM) [23–25], and many recent papers have developed the concept further [26–32]. The concept itself is very simple; two amplifiers are operated with constant envelope input signals (hence, very power-efficient), and their outputs are summed to produce the desired signal. The desired envelope and phase variation at the output is obtained by varying the relative phases between the two signals. This is shown schematically in Figure 1.16. The desired phase variation between the two amplifiers was originally obtained using analog techniques and now is more typically approached with digital techniques.

1.4.1 Historical Perspectives on Outphasing Power Amplifiers

A rereading of the original description of the concept from Chireix provides a very powerful motivation for the use of the outphasing power amplifier

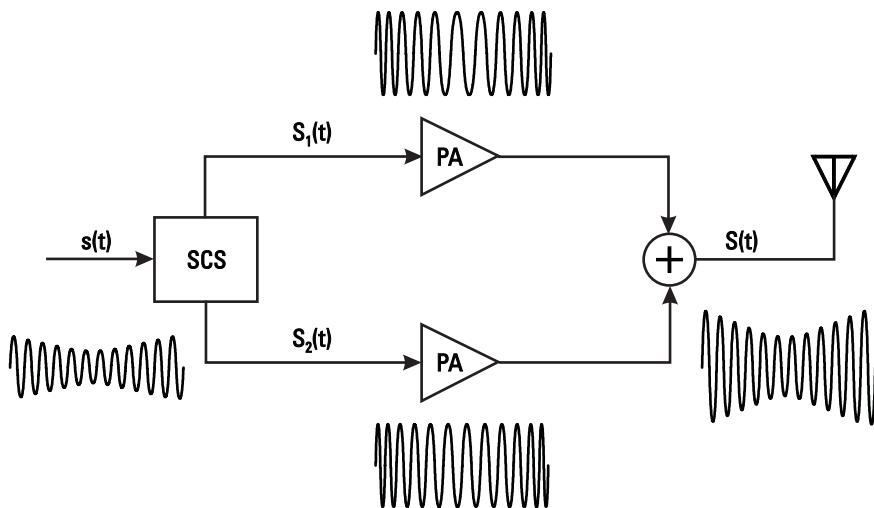


Figure 1.16 Simplified block diagram for an outphasing power amplifier.

technique, and shows that many of the problems that confront the designer of wireless power amplifiers today also confronted the early pioneers in the field. In their case, however, the challenges are applied to kilowatt-level AM transmitters! In fact, we will quote here directly from the original paper, whose introduction neatly summarizes the various power amplifier approaches.

All the present modulation systems can be classified in two categories; either the anode voltage of the high-frequency power stage is controlled by means of a low-frequency amplifier of suitable power in such a manner as to vary the output of the stage, or, the anode voltage remaining constant, the output power is controlled by an action on the grids of the high frequency power amplifier [19].

Here, Chireix is summarizing existing techniques for power amplification, either directly modulating the power supply with a low-frequency signal, or modulating the input with the AM carrier. Both techniques are still in widespread use today, albeit applied to transistor-based amplifiers rather than tube-based approaches. Both techniques have clear disadvantages, and these problems still exist today—nearly 70 years later. The paper then moves on to a proposed third approach:

Another path leading to the same results consists in dividing the amplifier into two parts each including one or several tubes according to the power required and each part having its own output circuit while the load is coupled differentially to both. The variable load is then obtained by acting on the phase difference between the grid excitation of the two parts of the final amplifier, whence the name of “outphasing” modulation given to the system [19].

This last excerpt provides a concise description of the outphasing power amplifier concept that is the basis for this book.

1.4.2 Introduction to the Theory of Outphasing Amplification

The basis for the outphasing amplifier concept is that an amplitude- and phase-modulated signal is resolved into two outphased *constant envelope* signals that are applied to highly efficient—and highly nonlinear—power amplifiers, whose outputs are summed. A block diagram of this approach is shown in Figure 1.16.

The separation of the bandpass signal is accomplished by the SCS. The detailed analysis of signal separation can be found in [26–28], and a brief mathematical description is given below. A complex representation of the band-limited source signal can be written as

$$s(t) = r(t)e^{j\theta(t)}; \quad 0 \leq r(t) \leq r_{\max} \quad (1.23)$$

This signal is split by the SCS into two signals with modulated phase and constant amplitudes, as illustrated in Figure 1.17.

$$S_1(t) = s(t) - e(t) \quad (1.24)$$

$$S_2(t) = s(t) + e(t) \quad (1.25)$$

where $e(t)$ is the quadrature signal and defined by

$$e(t) = js(t)\sqrt{\frac{r_{\max}^2}{r^2(t)} - 1} \quad (1.26)$$

The two signals are then amplified separately and sent to a power combiner. With the power combining, the in-phase signal components add

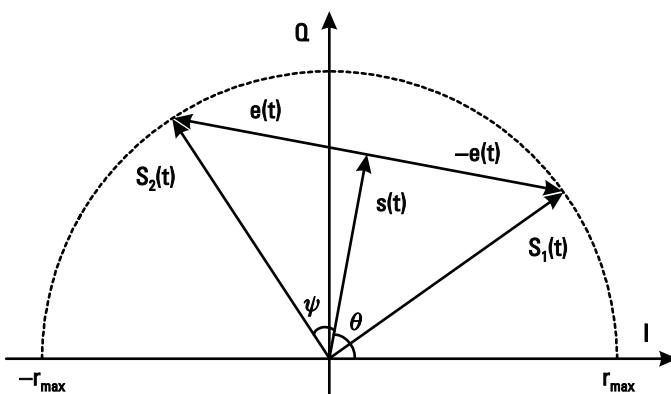


Figure 1.17 Separation of two component signals from the source signal.

together and the out-of-phase signal components cancel out; the resultant signal is the desired amplified replica of the original source signal.

This unique approach takes apparent advantage of the constant envelope feature of the two outphased signal components. First, the outphasing system is insensitive to the nonlinear characteristic of the two individual power amplifiers. Even though the power amplifiers are nonlinear, the final output can be highly linear and free of intermodulation—a key consideration for bandwidth efficient wireless communications.

Furthermore, each amplifier can be operated in a very power-efficient mode, and the overall efficiency of the system could be potentially high. In a word, the outphasing power amplifier is a promising approach that offers wideband high linearity and high efficiency simultaneously.

The above analysis is only valid when the characteristics of the two power amplifier branches are well-balanced. In practice, the matching conditions are difficult to achieve and maintain, due to process variation, thermal drift, component aging, and transition of channels. These mismatches can be expressed in terms of path imbalance, including gain and phase imbalance. In the presence of path imbalance, the resultant output is the sum of the desired signal and a distorted signal, due to incomplete cancellation of the quadrature signal $e(t)$. In contrast to the narrowband source signal, the spectrum of quadrature signal extends far into adjacent channels [33]. As a result, the incomplete cancellation of wideband components leaves a residue in adjacent channels, introducing adjacent channel interference (ACI) and alternate channel interference. Typical requirements for CDMA applications are on the order of a less than 0.3-degree phase mismatch and a less than 0.5-dB gain mismatch, a near impossibility in most

practical cases. Several different calibration schemes have been developed to minimize the effects of this limitation [33–38]. These techniques typically characterize the path imbalance through a simple feedback loop.

Frequency translation is frequently involved in order to upconvert the modulated baseband signal to the desired carrier frequency, and that leads to another important source of interference—the quadrature error of the in-phase/quadrature (I/Q) modulator. Quadrature error originates differently from the path imbalance, but leads to basically the same consequences on the output spectrum—both out-of-band interference and inband distortion. Efforts have been made to relate quadrature error to system linearity performance [39]. Due to its inherent randomized nature, quadrature error is difficult to compensate.

Over the years, the complexity of the analog SCS implementation has prevented the outphasing approach from being widely accepted. Today, the evolution of the DSP technique makes it possible to implement the SCS in software by using a standard DSP device. With this scheme, all processing is accomplished in baseband, and the baseband digital streams are converted to analog waveforms with a digital-to-analog converter (DAC) and reconstruction filter before frequency upconversion. The DSP brings about the third source of interference—quantization noise. The source signal and quadrature signal quantization errors result in different consequences [32].

A great advantage of the outphasing approach is that the two power amplifiers can operate at their peak efficiency, and their overall power efficiency is potentially high. The purely reactive power combiner achieves the highest power efficiency but presents a variable load impedance to the two power amplifiers depending on the required output power. This situation should be avoided in practice, since the load impedance presented to each amplifier appears highly reactive over a large portion of the cycle, harming the efficiency. A partial solution can be achieved using the so-called Chireix power-combining technique, illustrated in Figure 1.18. In this case, two quarter-wave transmission line impedance transformers and a shunt susceptance is added in order to improve the efficiency. The added shunt susceptance (which is inductive in one branch and capacitive in the other branch) cancels out the varying susceptance seen by each amplifier at one particular output power.

The microwave hybrid combiner is a convenient way of power combining to prevent the variable load impedance. However, the quadrature signal portion of power turns into waste heat. As a result, the amplifier only achieves its peak operating efficiency at maximum output power, and its

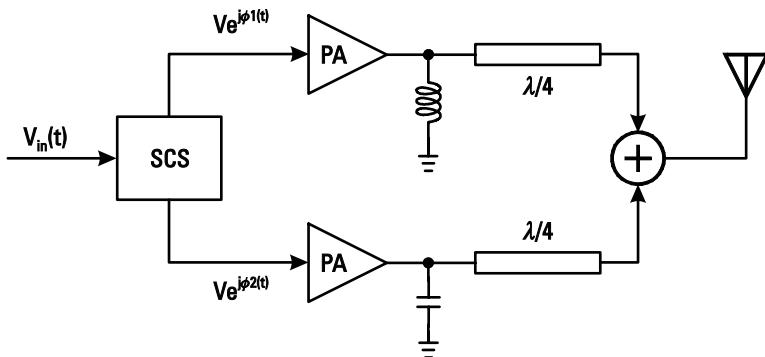


Figure 1.18 Simplified block diagram for an outphasing power amplifier showing transmission line coupling between the two power amplifiers.

efficiency decreases linearly as the output power decreases. This efficiency behavior is similar to that of a Class A amplifier, which is known to have a very poor overall efficiency. Of course, the peak efficiency of this outphasing approach is much higher than that of the Class A amplifier, but it would be desirable to do even better at low output powers.

References

- [1] Rapaport, T., *Wireless Communications, Principles and Practice*, Englewood Cliffs, NJ: Prentice Hall, 1996.
- [2] Couch, L. W., *Digital and Analog Communication Systems*, 5th ed., Upper Saddle River, NJ: Prentice Hall, 1997.
- [3] Feher, K., *Wireless Digital Communications: Modulation and Spread Spectrum Applications*, Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [4] Proakis, J., *Digital Communications*, 2nd. ed., New York: McGraw Hill, 1989.
- [5] Viterbi, A. J., *CDMA: Principles of Spread Spectrum Communication*, Reading, MA: Addison-Wesley Publishing Co., 1995.
- [6] Sevic, J., Notes on “Design of Power Amplifiers for Wireless Communications,” 1996.
- [7] Mehrotra, A., *GSM System Engineering*, Norwood, MA: Artech House, 1997.
- [8] EIA/TIA IS-95A, “Mobile Station—Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System,” Telecommunications Industry Association, 1995.
- [9] Agilent Application Note, *Characterizing Digitally Modulated Signals with CCDF Curves*, 2000.

- [10] Voelker, K. M., "Apply Error Vector Measurements in Communications Design," *Microwave and RF*, December 1995, pp. 143–152.
- [11] Heutmaker, M. S., "The Error Vector and Power Amplifier Distortion," in *Proc. 1997 Wireless Communications Conference*, Boulder, CO, 1997, pp. 100–104.
- [12] Agilent Application Note 1355, *Designing and Testing 3GPP WCDMA Transceiver Base Stations*, 2001.
- [13] Sevic, J., "Statistical Characterization of RF Power Amplifier Efficiency for IS-95 CDMA Digital Wireless Communication Systems," in *Proc. Second Annual Wireless Communications Conf.*, August 1997.
- [14] Mansell, A., and A. Bateman, "Adaptive Digital Predistortion Linearization," *Proc. Microwaves and RF Conference*, London, October 1996, pp. 270–275.
- [15] Cripps, S., *RF Power Amplifiers for Wireless Communications*, Norwood, MA: Artech House, 1999.
- [16] Petrovic, V., and C. Smith, "Reduction of Intermodulation Distortion by Means of Modulation Feedback," *IEE Conference on Radio Spectrum Conservation Techniques*, September 1983, pp. 44–49.
- [17] Seidel, H., "A Microwave Feed-Forward Experiment," *Bell System Technical Journal*, Vol. 50, November 1971, pp. 2879–2916.
- [18] Black, H. S., "Translating System," U.S. Patent 1,686,792, issued October 29, 1928.
- [19] Chireix, H., "High Power Outphasing Modulation," in *Proc. IRE*, Vol. 23, November 1935, pp. 1370–1392.
- [20] Cox, D. C., "Linear Amplification with Nonlinear Components," *IEEE Trans. Commun.*, Vol. COM-23, December 1974, pp. 1942–1945.
- [21] Cox, D. C., and R. P. Leck, "Component Signal Separation and Recombination for Linear Amplification with Nonlinear Components," *IEEE Trans. Commun.*, Vol. COM-23, November 1975, pp. 1281–1287.
- [22] Cox, D. C., and R. P. Leck, "A VHF Implementation of a LINC Amplifier," *IEEE Trans. Commun.*, Vol. COM-24, September 1976, pp. 1018–1022.
- [23] Bateman, A. "The Combined Analogue Locked Loop Universal Modulator (CALLUM)," *Vehicular Technology Society 42nd VTS Conf.*, May 1992, pp. 759–763.
- [24] Chan, K. Y., and A. Bateman, "Linear Modulators Based on RF Synthesis: Realization and Analysis," *IEEE Trans. Veh. Technol.*, Vol. 42, No. 6, June 1995, pp. 321–333.
- [25] Jennings, D. J., and J. P. McGeehan, "A High-Efficiency RF Transmitter Using VCO-Derived Synthesis: CALLUM," *IEEE Trans. Veh. Technol.*, Vol. 47, No. 6, June 1999, pp. 715–721.
- [26] Couch, L., and J. L. Walker, "A VHF LINC Amplifier," *Proc. IEEE Southeastcon'82*, Destin, FL, April 4–7, 1982, pp. 122–125.

- [27] Hetzel, S. A., A. Bateman, and J. P. McGeehan, "LINC Transmitter," *Elect. Lett.*, Vol. 27, No. 10, May 9, 1991, pp. 844–846.
- [28] Casadevall, F. J., and J. J. Olmos, "On the Behavior of the LINC Transmitter," *Proc. 40th IEEE Veh. Technol. Conf.*, Orlando, FL, May 6–9, 1990, pp. 29–34.
- [29] Raab, F. H., "Efficiency of Outphasing RF Power-Amplifier Systems," *IEEE Trans. Commun.*, Vol. COM-33, No. 10, October 1985, pp. 1094–1099.
- [30] Casadevall, F. J., and A. Valdovinos, "Performance Analysis of QAM Modulations Applied to the LINC Transmitter," *IEEE Trans. Veh. Technol.*, Vol. 42, November 1993, pp. 399–406.
- [31] Sundstrom, L., and M. Johansson, "Effect of Modulation Scheme on LINC Transmitter Power Efficiency," *Elect. Lett.*, Vol. 30, No. 20, September 29, 1994, pp. 1643–1645.
- [32] Sundstrom, L., "The Effect of Quantization in a Digital Signal Component Separator for LINC Transmitters," *IEEE Trans. Veh. Technol.*, Vol. 45, No. 2, May 1996, pp. 346–352.
- [33] Sundstrom, L., "Automatic Adjustment of Gain and Phase Imbalances in LINC Transmitters," *Elect. Lett.*, Vol. 31, No. 3, February 2, 1995, pp. 155–156.
- [34] Tomisato, S., K. Chiba, and K. Murota, "Phase Error Free LINC Modulator," *Elect. Lett.*, Vol. 25, No. 9, April 27, 1989, pp. 576–577.
- [35] Ampem-Darko, S., and H. S. Al-Raweshidy, "Gain/Phase Imbalance Cancellation Technique in LINC Transmitters," *Elect. Lett.*, Vol. 34, No. 22, October 1998, pp. 2093–2094.
- [36] Olson, S. A., and R. E. Stengel, "LINC Imbalance Correction Using Baseband Preconditioning," *1999 IEEE Radio and Wireless Conf. (RAWCON 99)*, Denver, CO, August 1999, pp. 179–182.
- [37] Zhang, X., and L. E. Larson, "Gain and Phase Error Free LINC Transmitter," *IEEE Trans. Veh. Technol.*, Vol. 49, No. 5, September 2000, pp. 1986–1994.
- [38] Zhang, X., L. E. Larson, and P. M. Asbeck, "Calibration Scheme for LINC Transmitter," *Elect. Lett.*, Vol. 37, No. 5, March 2001, pp. 317–318.
- [39] Sundstrom, L., "Spectral Sensitivity of LINC Transmitters to Quadrature Modulator Misalignments," *IEEE Trans. Veh. Technol.*, Vol. 49, No. 4, July 2000, pp. 1474–1487.

2

Linearity Performance of Outphasing Power Amplifier Systems

2.1 Introduction

As discussed in Chapter 1, in an outphasing power amplifier system, the AM of a bandpass waveform is transformed into PM of two signals prior to power amplification, and finally recovered by recombining the two amplified signals. As shown in Figure 1.16, with this approach, two highly efficient nonlinear power amplifiers can be used without distorting the original signal. This is the main advantage of the outphasing approach to linear power amplification. However, the bandlimited feature is no longer preserved—a wideband signal is added to and subtracted from the original source input to construct the two outphased component signals. Since the spectrum of the two additional signal components extends far into the adjacent channels, the system linearity performance relies on the precise cancellation of the wideband signals during power combining, and the tolerance on the matching between the two power amplifier branches becomes critical—a slight mismatch between the two paths can result in significant distortion in the adjacent and alternate frequency bands.

The gain and phase characteristics of the two amplifier branches, which includes the digital baseband, analog, and RF portions, must be well-matched for the acceptably small out-of-band rejection. A failure to achieve sufficient matching generally creates both out-of-band interference and inband distortion. The out-of-band interference is characterized by the ACPR, while the inband distortion may be described in terms of EVM. Various factors can contribute to path mismatch and the consequent linearity

degradation. These include environmental variations such as thermal drift, component aging, and channel transition. A number of coupling effects in integrated circuits—substrate, capacitive, bond wire, and package—as well as process variation during fabrication could also contribute to a significant amount of mismatch. The main contributors to the mismatch are the RF gain and phase errors, the quadrature modulator errors, and the imbalance caused by the SCS. If the SCS is implemented with DSP, the effects of the quantization noise, DSP sampling rate, and reconstruction filter also need to be investigated.

This chapter begins with a brief discussion of the relevant modulation schemes and baseband filtering typically used in mobile communications applications. Next, various techniques for the implementation of the signal component separation will be described. The major sources of imbalance, including the path imbalance, the quadrature modulator error, the SCS quantization, the DSP sampling rate, and reconstruction filtering, are successively investigated. Finally, we summarize some of the practical implementation issues for outphasing power amplifiers.

2.2 Digital Modulation Techniques

Modulation is a process that encodes source information onto a carrier signal for optimized transmission. During the modulation process, the analog or digital source signal is mapped onto the amplitude or phase (or both) of the carrier signal. Digital modulation offers a number of advantages over its analog counterpart, such as increased capacity, tolerance to channel impairments and noise, and accommodation to various digital signal conditioning and processing techniques [1–4]. Various modulation schemes and their variants have been developed. Their names are usually derived from the mapping operation, such as PSK, FSK, and quadrature amplitude modulation (QAM). This section briefly introduces several digital modulation techniques that are used in modern digital communications, including QPSK, OQPSK, $\pi/4$ -DQPSK, and QAM. GMSK will be discussed in Section 2.3, together with the Gaussian filter.

2.2.1 QPSK and Its Variations

QPSK belongs to the family of M-ary phase shift keying modulation, in which the digital data is mapped onto M-number of discrete phase states. As shown in Figure 2.1(a), four phase states of $\pm 45^\circ$ and $\pm 135^\circ$ are used in QPSK, and

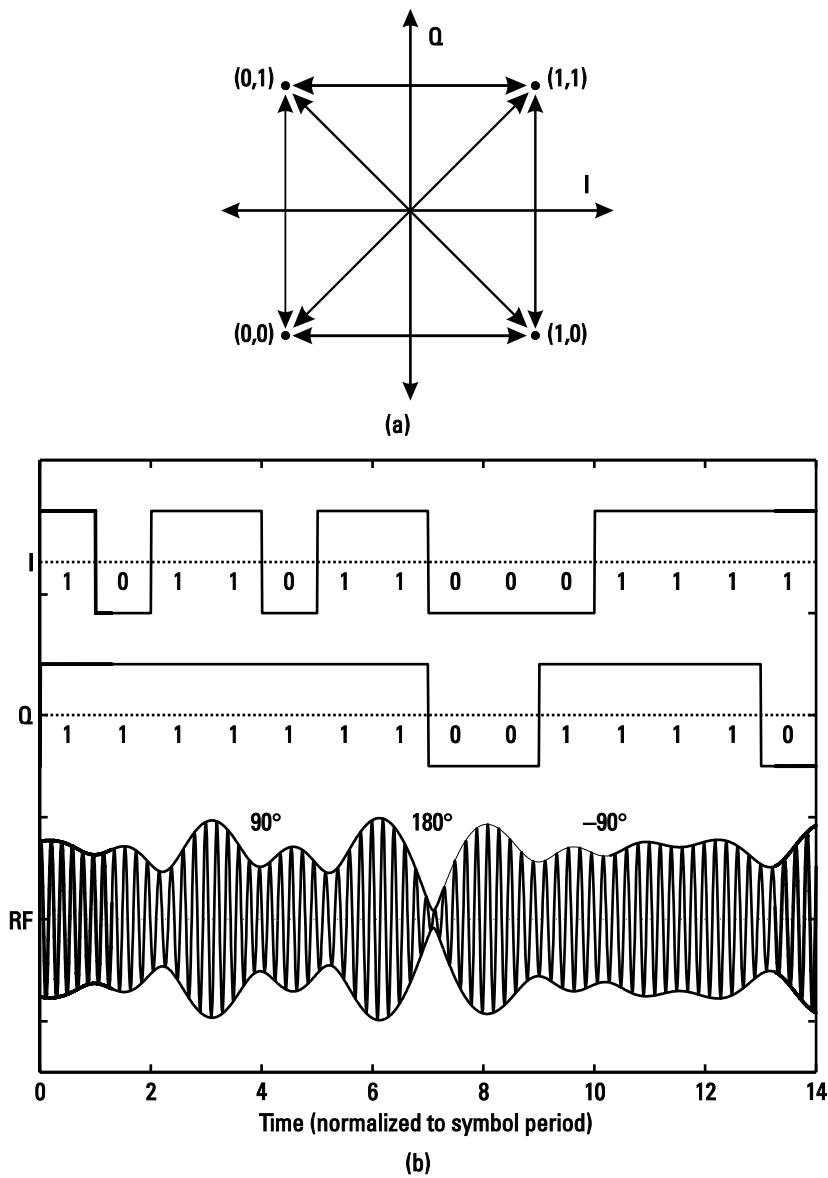


Figure 2.1 (a) QPSK signal constellation and phase transitions, and (b) a baseband-filtered time domain waveform indicating phase transition and envelope variation.

each phase state represents a single symbol (or two bits) of information. QPSK modulation can be realized by a quadrature modulator with the baseband I/Q data streams as the driving signals. Figure 2.1(a) also shows that the possible

phase transitions are -90° , $+90^\circ$, or 180° . During a 180° transition, the signal trajectory passes through the origin, which is also known as a “zero-crossing.” The zero-crossing leads to the well-known problem of spectral regrowth, when the QPSK waveform passes through the baseband filter and is processed by a saturated power amplifier. Baseband filtering, which is discussed in Section 2.3, is used to suppress the sidelobe spectrum, improving the spectral efficiency and minimizing the interference spreading to adjacent channels. However, it introduces a variation on the envelope of the filtered QPSK waveform. Specifically, every time a zero-crossing is encountered, the envelope of the filtered signal drops to zero. This situation is illustrated in Figure 2.1(b). The baseband filter used is a square-root raised cosine filter with roll-off 0.35, described in Section 2.3. A saturated power amplifier possesses high power efficiency but usually exhibits significant nonlinearity when confronted with a high degree of variation in the envelope of the input signal. Power amplifier nonlinearity results in a spectral leakage to the neighboring channels, excited by the signal envelope fluctuation. Hence, QPSK-modulated waveforms generally require a highly linear power amplifier to accommodate the substantial envelope variation.

The zero-crossing may be eliminated if an appropriate time offset is introduced between the I/Q data streams. In OQPSK, the Q-channel data stream is delayed by half the symbol period and the phase transitions are therefore restricted to $\pm 90^\circ$. Figure 2.2(a) illustrates the signal constellation and the phase transitions in OQPSK. As shown in Figure 2.2(b), the filtered OQPSK waveform exhibits much less envelope fluctuation and is therefore much more attractive for linear power amplifier implementation. Due to the similarity, OQPSK has the same BER and power spectral density (PSD) as QPSK. OQPSK is used in the CDMA IS-95 reverse link.

The $\pi/4$ -DQPSK modulation technique is another variation of QPSK that differs from the latter in two aspects. First, the data stream is differentially encoded such that the information bits are embedded in the phase change between the successive symbols rather than the absolute phase. As an example, in differential binary PSK (DBPSK) modulation, a digital “1” indicates a 180° change on the absolute phase of the carrier, while a “0” results in no phase change, or vice versa. One advantage with $\pi/4$ -DQPSK is that the differential encoding enables “noncoherent” detection, which simplifies the receiver implementation and is especially desirable for mobile communications. Second, an additional 45° phase shift is inserted in each symbol such that there exist eight possible phase states while the phase transition is restricted to $\pm 45^\circ$ and $\pm 135^\circ$. The constellation and phase transition of $\pi/4$ -DQPSK is illustrated in Figure 2.3(a). This scheme is

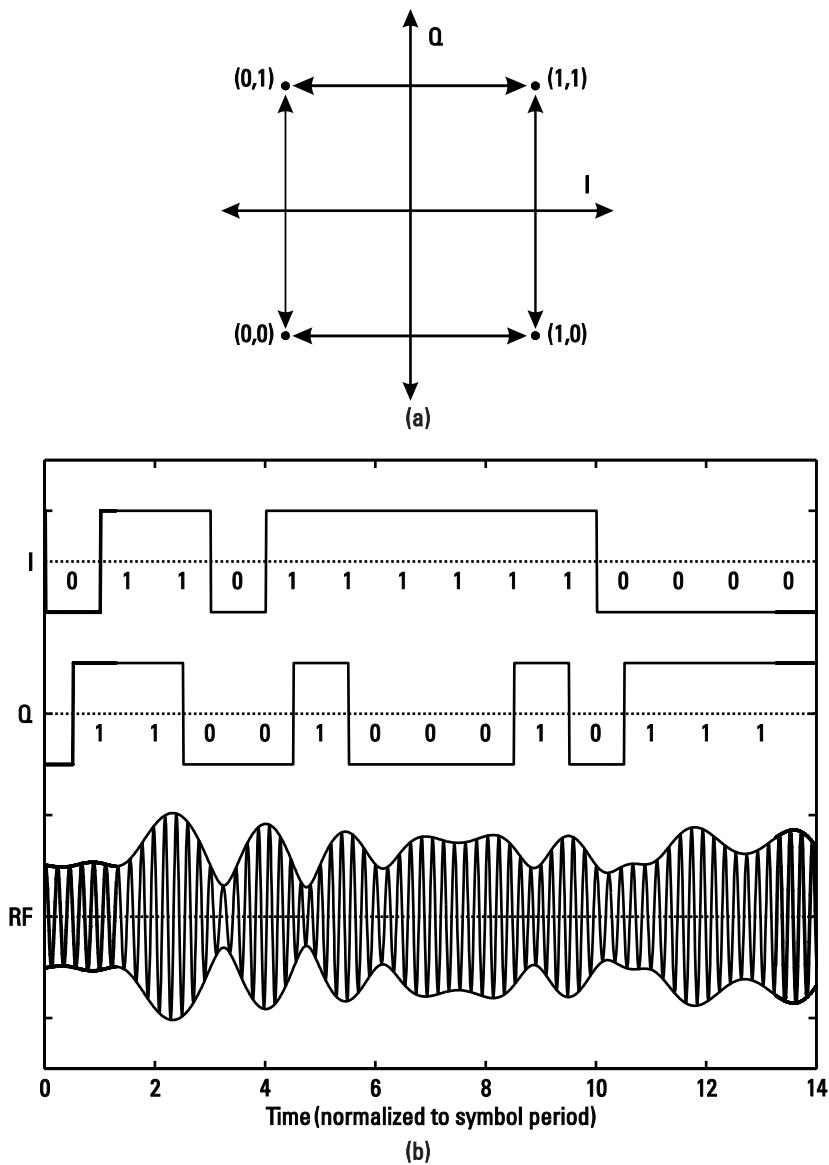


Figure 2.2 (a) OQPSK signal constellation and phase transitions, and (b) a baseband-filtered time domain waveform indicating phase transition and envelope variation.

therefore a compromise between QPSK and OQPSK, and the maximum phase transition of 135° leads to less envelope variation compared to QPSK, as shown in Figure 2.3(b). The 45° phase shift also ensures the continuous

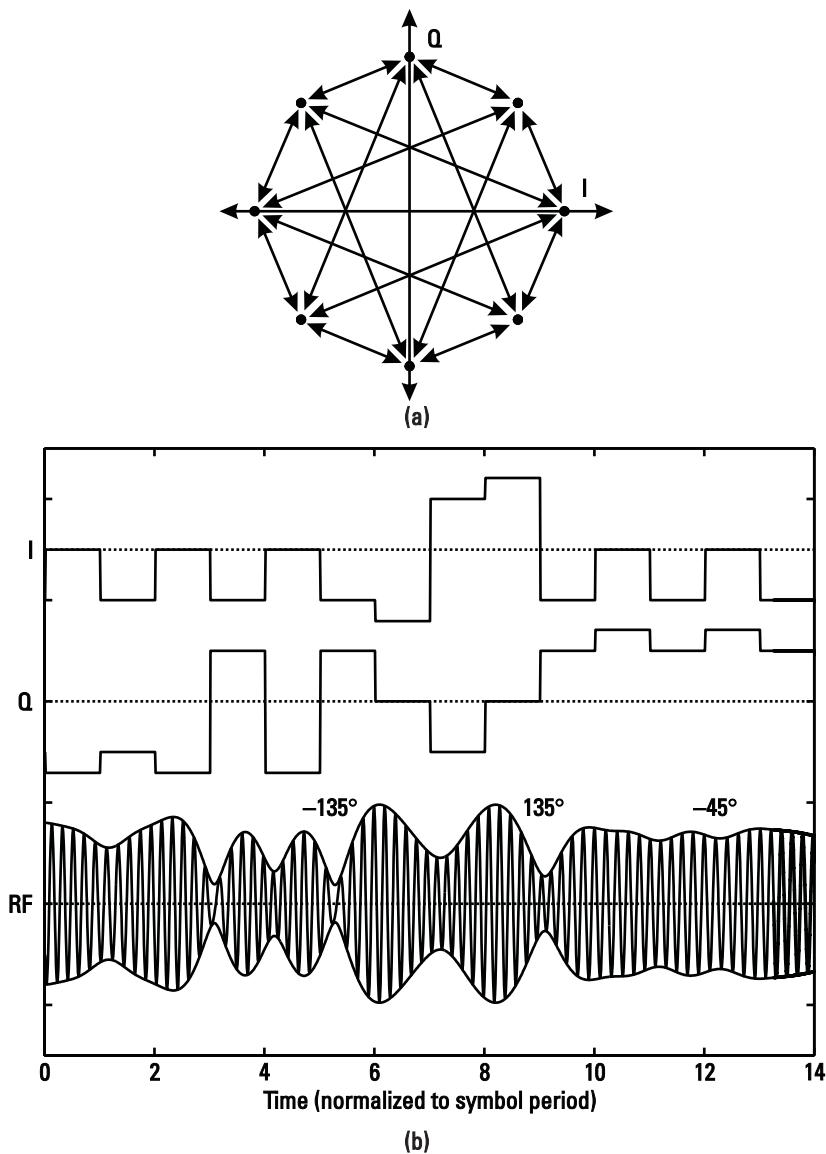


Figure 2.3 (a) $\pi/4$ -DQPSK signal constellation and phase transitions, and (b) a baseband-filtered time domain waveform indicating phase transition and envelope variation. I and Q data are encoded.

phase transition for each symbol that allows the proper operation of the timing recovery and synchronization circuits in the receiver. One disadvantage of $\pi/4$ -DQPSK is the higher BER compared to QPSK.

It should be pointed out that a higher envelope variation of the bandlimited signal, or more precisely, a higher PAP, does not necessarily lead to greater spectral regrowth or intermodulation [5]—for example, when the signal is fed to a power amplifier with modest nonlinearities, like a Class AB or B design. Conventionally, the power amplifier is backed off from its 1-dB compression point approximately by an amount of PAP to accommodate the signal envelope variation. The intermodulation is generated during power amplification when the high input level drives the amplifier near saturation. Therefore, the average amount of intermodulation generated by the power amplifier is dependent on the statistical behavior of the instantaneous signal power above the average point. This time-domain distribution is characterized by the CCDF or envelope distribution function. From [5], although the QPSK-modulated signal possesses a higher value of PAP, it actually exhibits less spectral regrowth than OQPSK- and $\pi/4$ -DQPSK-modulated signals on the same average power basis. The power amplifier used was a Class AB, and the filters were CDMA IS-95 or IS-54 baseband filters. However, a higher PAP does result in a lower power efficiency in many cases.

2.2.2 QAM

QAM is a modulation technique in which the symbols have both amplitude and phase variations, as compared to PSK modulation techniques where only the phase of the carrier is varied. A M-ary QAM modulation combines every $\log_2 M$ bit as an individual symbol, and the symbol rate is hence $\log_2 M$ times less than the bit data rate. A typical value for M is in the range of 16–1,024. QAM is a spectrally efficient modulation scheme, and, it is used by computer modems and wireless LAN systems. Figure 2.4 illustrates the constellation of 16-QAM, with each symbol representing four bits. In some applications, the I/Q data streams are offset by half the symbol period, just as with OQPSK.

2.3 Baseband Filtering of Digital Data

The transmission of purely digital data across the wireless communication channel represents an interesting problem for those concerned with the spectral efficiency of the communications system. In this case, the spectral efficiency of a digital signal is defined as the number of bits per

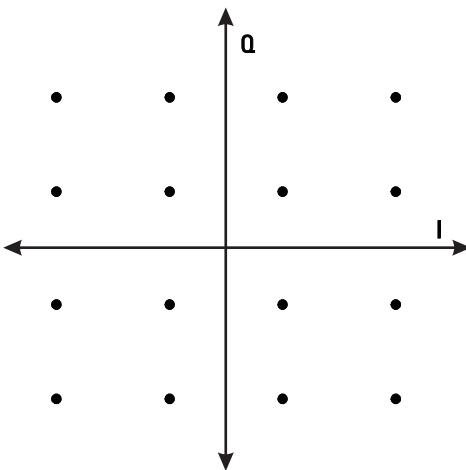


Figure 2.4 16-QAM constellation.

second of data that can be supported by each hertz of channel bandwidth, that is,

$$\eta = \frac{R}{B} \quad (2.1)$$

where R is the data rate, and B is the channel bandwidth. Clearly, a higher value of η is more desirable than a lower value. Intuitively, as the noise introduced by the channel grows, the maximum value of η will decline. The maximum spectral efficiency is given by Shannon's channel capacity formula [6]

$$\eta_{\max} = \ln\left(1 + \frac{S}{N}\right) \quad (2.2)$$

where S/N is the signal-to-noise ratio (SNR) of the received signal. A delightful “proof” of this theorem can be found in [7]. Unfortunately, the total bandwidth of a square-wave digital signal is infinite, containing spectral energy at all of the harmonics of the fundamental frequency. A polar signal (a digital sequence consisting of $+1$ or -1) consisting of a sequence of random binary data has a spectral response of

$$P_{\text{polar}}(f) = A^2 T_b \frac{\sin^2(\pi f T_b)}{(\pi f T_b)^2} \quad (2.3)$$

and a unipolar signal (a digital sequence consisting of +1 or 0) consisting of a sequence of random binary data has a spectral response of

$$P_{\text{unipolar}}(f) = \frac{A^2 T_b}{4} \frac{\sin^2(\pi f T_b)}{(\pi f T_b)^2} \left[1 + \frac{1}{T_b} \delta(f) \right] \quad (2.4)$$

where A is the amplitude of each bit, T_b is the time period associated with each bit, and $\delta(f)$ is the Dirac function. Clearly, this function has significant energy to a frequency much higher than $1/T_b$. The one-sided spectral densities for these two cases are shown in Figure 2.5. Therefore, a crucial distinction between analog and digital transmission is the requirement for filtering the data in the case of digital transmission to minimize the required bandwidth.

As a result, the digital data is typically filtered prior to modulation and upconversion. This filtering improves the spectral efficiency dramatically, but at the potential cost of another phenomena known as intersymbol interference (ISI). The process of lowpass filtering the digital data narrows its response in the frequency domain, but because of the inverse relation

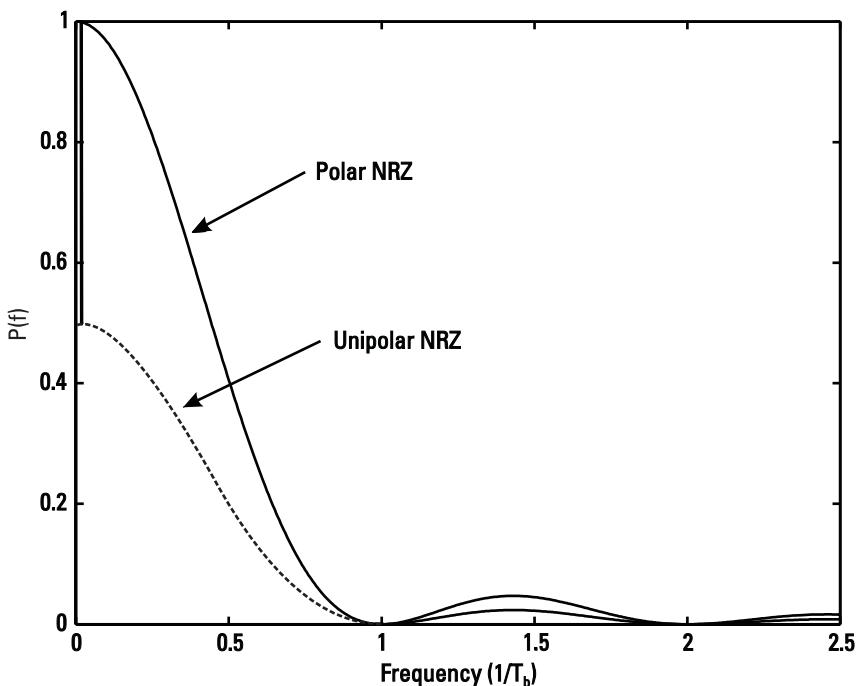


Figure 2.5 Spectral density of unipolar NRZ data and polar NRZ data.

between the time and frequency domain characteristics of any signal (a narrow signal in the time domain will create a broad signal in the frequency domain, and vice versa), the time response of each pulse is widened. This is often referred to as the uncertainty principle in signal processing—analogous in some ways to Heisenberg's uncertainty principle in physics. For example, the RC lowpass filter of Figure 2.6 illustrates a random stream of “1’s” and “0’s” passing through the filter. Note that each bit is potentially corrupted by the long “tails” of the responses of the previous bits, and ISI results. If the data was perfectly periodic, then the system would settle down into a steady state response and there would be no ISI, but of course there would be no information transmitted in such a case.

There are several possible solutions to this problem. The most commonly used is known as Nyquist's first method, where the impulse

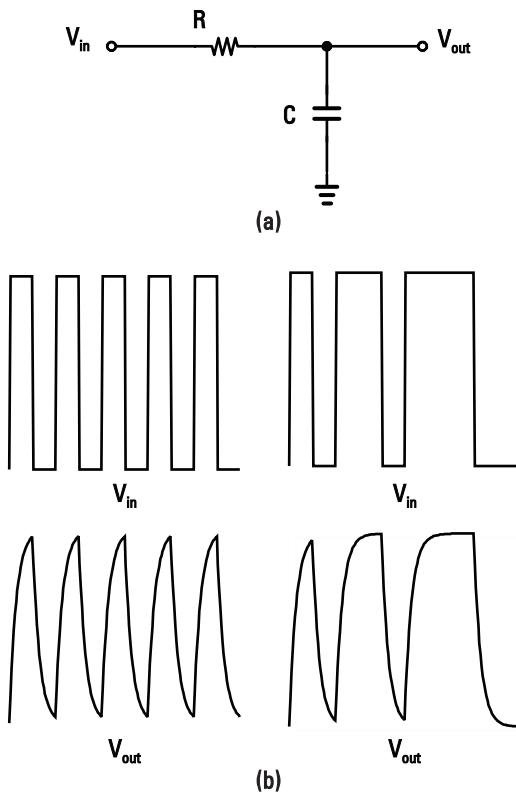


Figure 2.6 ISI due to lowpass filtering: (a) lowpass RC filter, and (b) digital data passed through lowpass filter exhibits ISI.

response of the filter used to improve the bandwidth efficiency of the system has the property that

$$h(kT_b + \tau) = \begin{cases} C, & k = 0 \\ 0, & \text{otherwise} \end{cases} \quad (2.5)$$

This feature of the filter eliminates ISI exclusively at times that are integral multiples of the bit period. It is of course impossible to eliminate ISI at all times, but if it can be eliminated at the instant of sampling, then the problem can be dramatically reduced. The class of filters that satisfy this criteria are known as Nyquist filters. One example of an “ideal” Nyquist filter exhibits a *sinc* ($\sin x/x$) behavior for its impulse response, such that

$$h(t) = \frac{\sin\left(\pi \frac{t}{T_b}\right)}{\pi \frac{t}{T_b}} \quad (2.6)$$

Note that the impulse response of this filter is zero for all values of $t = nT_b$, satisfying Nyquist’s criterion. The filter also has the desirable property that it behaves like a “brick-wall” filter in the frequency domain. However, this ideal filter is physically unrealizable, since its impulse response is noncausal. In addition, its peak amplitude decays only as $(1/t)$, and so any mistiming in the clock recovery circuit can generate significant ISI. One of the filters that satisfy Nyquist’s criterion is the *raised cosine* filter family. The raised cosine filter has an impulse response that drops off much faster in time than the *sinc* response, at the expense of a small increasing bandwidth, and hence finds its applications in wireless communications.

2.3.1 Raised Cosine Filter

The raised cosine filter derives its name from its shape in the frequency domain. The frequency response of a raised cosine filter can be described by

$$H(f) = \begin{cases} T_s & 0 \leq |f| \leq \frac{1-\alpha}{2T_s} \\ \frac{T_s}{2} \left\{ 1 + \cos \left[\frac{\pi T_s}{\alpha} \left(|f| - \frac{1-\alpha}{2T_s} \right) \right] \right\} & \frac{1-\alpha}{2T_s} \leq |f| \leq \frac{1+\alpha}{2T_s} \\ 0 & |f| > \frac{1+\alpha}{2T_s} \end{cases} \quad (2.7)$$

where α is the roll-off factor that determines the excess bandwidth of the filter frequency response and in the range of [0,1], and T_s is the sampling or the symbol period. A zero roll-off is essentially a *sinc* response. As shown in Figure 2.7(a), the filter response has a cosine roll-off shape between $(1 - \alpha)/2 T_s$ and $(1 + \alpha)/2 T_s$ and becomes strictly zero beyond $(1 + \alpha)/2 T_s$. This frequency is thus the bandwidth of the raised cosine response

$$\text{BW} = \frac{1 + \alpha}{2} f_s \quad (2.8)$$

where $f_s = 1/T_s$ is the sampling rate or the symbol rate. The corresponding filter impulse response is described by

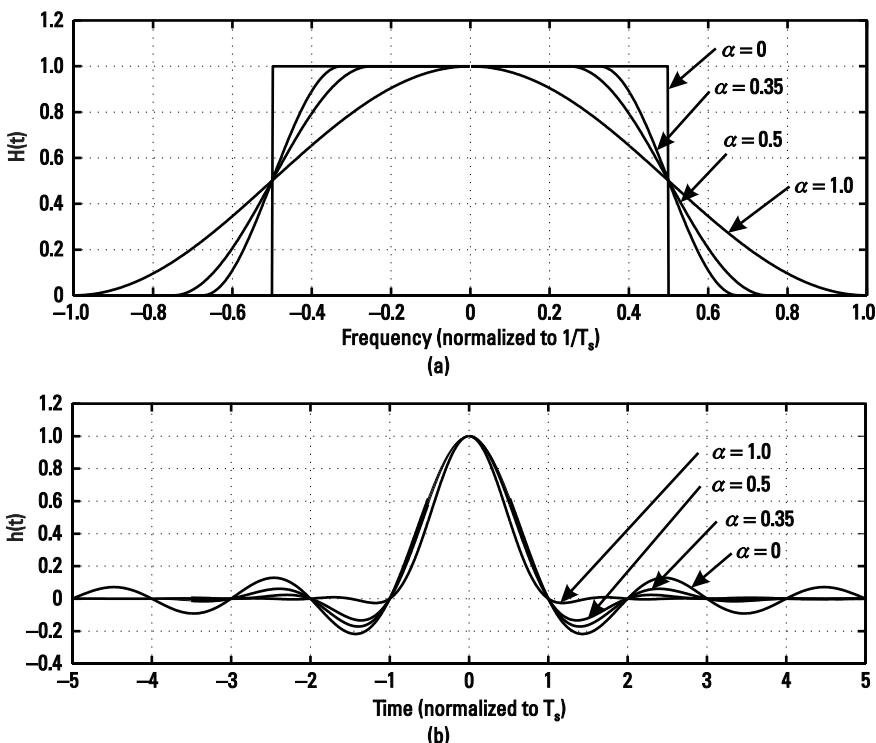


Figure 2.7 (a) Raised cosine filter frequency response, and (b) impulse response for various roll-off factors.

$$h(t) = \frac{\cos\left(\pi\alpha\frac{t}{T_s}\right)}{1 - \left(2\alpha\frac{t}{T_s}\right)^2} \cdot \frac{\sin\left(\pi\frac{t}{T_s}\right)}{\left(\pi\frac{t}{T_s}\right)} \quad (2.9)$$

It is clear that at the sampling instant $t = kT_s$, the impulse response is essentially zero except at $k = 0$. Therefore no ISI will be created. The peak amplitude of the raised cosine impulse response, with a modest value of α , drops off much faster than the *sinc* response. This is due to a more gradual roll-off of cosine shaping at the filter cutoff frequency. Note that the smaller the roll-off factor, the less the absolute bandwidth but the higher the impulse response ripple, as shown in Figure 2.7(b). The raised cosine filtering is widely used in several communication standards, and a value of α of 0.35 is used in IS-54, 0.5 in PDC, and 0.22 in WCDMA.

In most communications applications, the raised cosine filter is divided into two sections, one placed in the transmitter and the other in the receiver—each one is the so-called square-root raised cosine filter. With this arrangement, the two sections together form the Nyquist filter, while the filter in the receiver side also serves as a matched filter. The matched filter has a response that passes the desired signal, while it rejects the noise and interference and thus maximizes the SNR. In the presence of white noise, the matched filter has an impulse response that is simply the time-reversed transmitted signal pulse shape. The frequency response of a square-root raised cosine filter is the square root of (2.7) and the impulse response is given by

$$h_s(t) = \frac{4\alpha}{\pi\sqrt{T_s}} \frac{\cos\left[(1 + \alpha)\pi\frac{t}{T_s}\right] + \frac{\sin\left[(1 - \alpha)\pi\frac{t}{T_s}\right]}{\left(4\alpha\frac{t}{T_s}\right)}}{1 - \left(4\alpha\frac{t}{T_s}\right)^2} \quad (2.10)$$

The frequency and impulse responses are shown in Figure 2.8. The noise bandwidth of the square-root raised cosine filter is readily calculated with (2.7)

$$\begin{aligned} B &= \frac{1}{|H_s(0)|^2} \int_0^\infty |H_s(f)|^2 df \\ &= \frac{1}{2} f_s \end{aligned} \quad (2.11)$$

which is simply half the symbol rate.

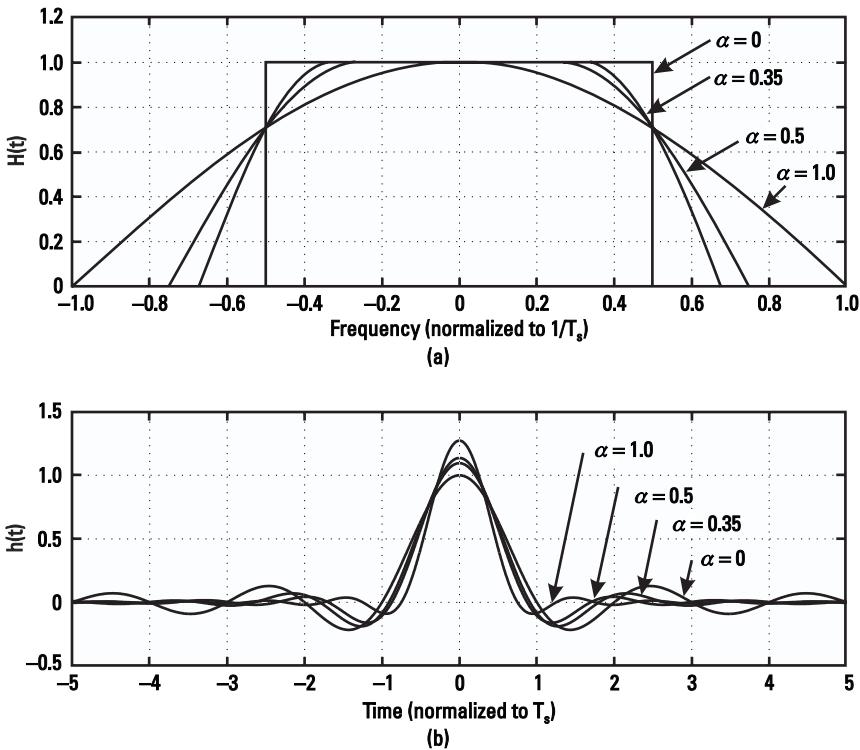


Figure 2.8 (a) Square-root raised cosine filter frequency response, and (b) impulse response for various roll-off factors.

The impulse response of the raised cosine filter lasts indefinitely and needs to be truncated for use in practical applications. The truncation results in a set of nonzero side lobes in the frequency domain. However, since the pulse response decay is proportional to $1/t^3$, a proper truncation results in little deviation from the theoretical performance. A smaller roll-off factor α results in a higher ripple in the impulse response and thus requires a longer pulse shape to achieve the same stopband attenuation. For a stopband attenuation of approximately 40 dB, a rough estimate on the required truncation length for $0.2 \leq \alpha \leq 0.75$ is given by

$$L_{\min} = -44\alpha + 33 \quad (2.12)$$

where L_{\min} is the truncated filter impulse response length.

The raised cosine filter may be implemented within the baseband DSP or with an analog filter. With the baseband DSP approach, the sampled

values of the filter shape are stored in a lookup table (LUT) and read out in sequence to form the desired symbols. The advantage of the analog approach to the realization of the filter is the reduced hardware complexity, since the raised cosine filter also serves as a reconstruction filter following the D/A converter.

2.3.2 Gaussian Filter

A Gaussian filter is used in GMSK, which is the modulation approach adopted for the GSM standard [8]. GMSK modulation is minimum phase shift keying (MSK) with Gaussian filtering and can be implemented by passing the nonreturn-to-zero (NRZ) rectangular waveform (which has values of ± 1) through a Gaussian filter followed by a frequency modulator with a modulation index of 0.5. GMSK is intrinsically a constant envelope modulation and thus can be applied directly to the nonlinear power amplifier without distortion. The frequency response of a Gaussian filter is described by

$$H(f) = e^{-(f/\alpha)^2} \quad (2.13)$$

where the parameter α is a constant that determines the filter 3-dB bandwidth B

$$B = \sqrt{\ln \sqrt{2}} \cdot \alpha = 0.5887\alpha \quad (2.14)$$

The impulse response of the Gaussian filter is

$$h(t) = \sqrt{\pi} \alpha e^{-(\pi \alpha t)^2} \quad (2.15)$$

Note that both the frequency response and impulse response have the same Gaussian shape. Gaussian filtering is usually specified by the relative bandwidth in terms of the bit rate of the data stream, for example,

$$BT = \frac{B}{f_b} \quad (2.16)$$

where f_b is the bit data rate. The fact that the Gaussian impulse response exhibits a smooth roll-off with no overshoot or ringing leads to little ISI. Gaussian filtering also suppresses the high-frequency components and

provides a narrower bandwidth compared to MSK modulation without filtering. As BT decreases, however, the filtered pulse tends to span over several symbol periods and more ISI is generated. GSM uses a relative bandwidth of 0.3 as a compromise between the transmission bandwidth and ISI.

2.3.3 IS-95 Baseband Filter

The CDMA IS-95 system uses a pulse-shaping filter that does not satisfy Nyquist's criterion and is included here for completeness. The IS-95 baseband filter uses a Chebyshev roll-off instead of cosine roll-off. Figure 2.9 shows the limit of the normalized frequency response of the baseband filter [9]. The passband edge frequency f_p is 590 kHz, and the stopband edge frequency f_s is 740 kHz. The passband ripples are bounded within ± 1.5 dB, and the stopband attenuation is equal to or greater than 40 dB. The impulse response of the baseband filter $h_c(t)$ satisfies the following equation

$$\sum_{k=0}^{\infty} [\alpha h_c(kT_s - \tau) - h(k)]^2 \leq 0.03 \quad (2.17)$$

where the constants α and τ are used to minimize the mean square error. The constant T_s is one-quarter of a PN chip. The values of the coefficients $h(k)$ are listed in [9] for $k < 48$. These values can be directly used as a 48-tap baseband FIR filter with 4× oversampling rate.

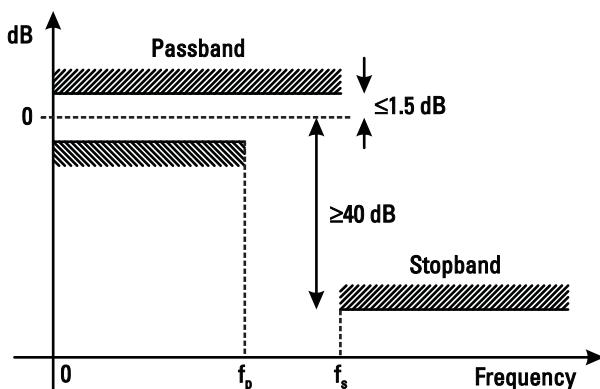


Figure 2.9 CDMA IS-95 baseband filter frequency response limits [9].

2.4 Signal Component Separation for Outphasing Amplifiers

Despite the neatness and elegance of the outphasing concept, the practical complexity of realizing the signal component separation has limited its wide acceptance. The difficulty stems from the fact that the generation of the two component signals involves memoryless nonlinear signal processing that requires a high degree of accuracy. During the past decades, various approaches have been proposed to accomplish this function, but difficulties remain in terms of such factors as implementation complexity, bandwidth, and power consumption, due to the stringent distortion and noise requirements.

There are two canonical forms for the signal component separation process—the phase modulation method and the in-phase/quadrature method [10]. In the phase modulation method, the phase modulation $\theta(t)$ and amplitude modulation $r(t)$ of the input are separated first, and two nonlinear phase modulators with inverse sine characteristics are used to transform the input envelope variation into the phase of the desired output signals, as shown in Figure 2.10. The two component signals can be described by [11–13]

$$S_1(t) = A_c \sin[\omega_c t + \theta(t) + \phi(t)] \quad (2.18)$$

$$S_2(t) = A_c \sin[\omega_c t + \theta(t) - \phi(t)] \quad (2.19)$$

where A_c is the amplitude of the carrier signal, and $0 \leq A_c \leq r(t)$. $\phi(t)$ is given by

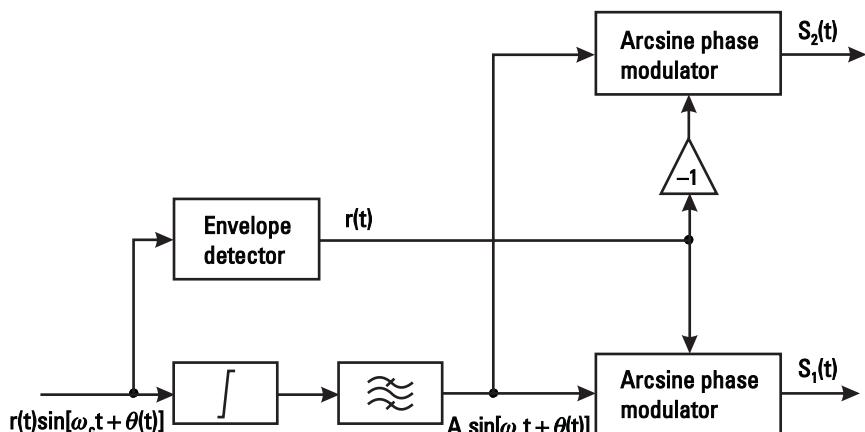


Figure 2.10 Signal component separation with nonlinear phase modulators [10].

$$\sin \phi(t) = \frac{r(t)}{A_c} \quad (2.20)$$

The desired output signal is obtained by subtracting $S_2(t)$ from $S_1(t)$ following the power amplification,

$$\begin{aligned} S(t) &= GS_1(t) - GS_2(t) \\ &= 2Gr(t) \cos[\omega_c t + \theta(t)] \end{aligned} \quad (2.21)$$

The SCS in this scheme is typically realized as an analog circuit and operates at some intermediate frequency or directly at the carrier frequency. Figure 2.11 shows an implementation of the inverse sine phase modulator, which is constructed by a linear phase modulator with the necessary feedback [11]. The operation of the circuit can be understood by making the following idealization—the input impedance of the amplifier G is much higher than R_f or R_i , so that the current injected into the input of the amplifier is negligible, and the loop gain is sufficiently large so that the amplifier input node V_i is close to a virtual ground. These assumptions give rise to the following equation

$$\frac{r(t)}{R_i} + \frac{V_f(t)}{R_f} = 0 \quad (2.22)$$

The phase modulator modulates the phase of the carrier signal by an amount linearly proportional to the amplifier output V_o ,

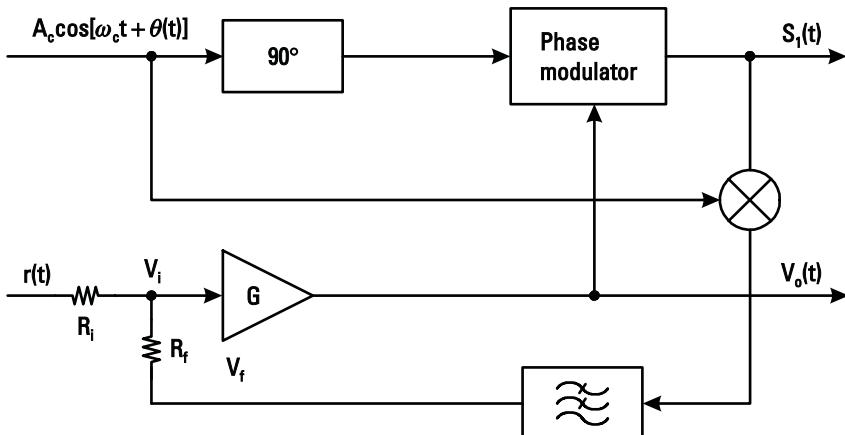


Figure 2.11 Nonlinear inverse sine phase modulator [11].

$$S_1(t) = A_c \sin[\omega_c t + \theta(t) + mV_o(t)] \quad (2.23)$$

where m is the characteristic of the phase modulator and A_c is the amplitude of the carrier signal. The mixer and lowpass filter act as a phase detector that downconverts the modulator output and removes the original phase modulation term $\theta(t)$, then

$$V_f(t) = \frac{1}{2} A_c^2 \sin[mV_o(t)] \quad (2.24)$$

Substituting (2.22) into (2.24), we have

$$mV_o(t) = -\arcsin\left[\frac{2R_i}{A_c^2 R_f} r(t)\right] \quad (2.25)$$

The amplifier output V_o can be made proportional to the inverse sine of the input envelope by adjusting the values of A_c , R_f , and R_i . As a bonus, the phase modulator output $S_1(t)$ is one of the constant-envelope vector components with the desired phase modulation. A similar circuit can generate the other component signal $S_2(t)$ by inverting the sign of the amplifier output $V_o(t)$. The nonideality of the limiter and envelope detector, the finite loop gain, and the loop noise all contribute to the performance degradation. Another limitation is that the feedback loop delay limits the bandwidth of the modulator.

As described in Chapter 1, the in-phase/quadrature method first computes the I/Q components of the quadrature signal $e(t)$,

$$e(t) = js(t)\sqrt{\frac{r_{\max}^2}{r^2(t)} - 1} \quad (2.26)$$

then the two component signals are obtained by summing and subtracting the quadrature signal from the source input.

$$S_1(t) = s(t) - e(t) \quad (2.27)$$

$$S_2(t) = s(t) + e(t) \quad (2.28)$$

A graphic illustration of the I/Q component separation is given in Figure 1.17.

The calculation of the quadrature signal involves complicated functions, such as multiplication, division, and square root, which are generally

difficult to be accurately and cost-effectively implemented. However, the in-phase/quadrature SCS can take the advantage of baseband signal processing, and two quadrature modulators are used to translate the component signals to the desired carrier frequency. Figure 2.12 illustrates an early version of an analog signal component separator that may operate at intermediate frequency (IF) or RF [10].

Another means to realize linear amplification with nonlinear power amplifiers with the outphasing approach is based on a feedback tracking loop to lock the power amplifier output to the input waveforms [14–17]. These approaches can be regarded as the generalization of the feedback power amplifier applied to the outphasing system. Two representative methods that fall into this category are CALLUM [14] and the vector-locked loop (VLL) [15], in which the amplifier output is compared to the source input and the error signal is used to control the relative phases of two VCOs. The former technique (CALLUM) is realized in Cartesian form while the latter (VLL) is realized in polar form. Since no explicit SCS is involved, and the feedback loop automatically corrects the path imbalance and improves the system linearity, these two structures will be discussed in Chapter 3.

The in-phase/quadrature method is most common and practical means for SCS, whether implemented in analog or digital circuitry. The rest of this section reviews the different ways to generate the quadrature signal, or in a sense, realize the inverse sine or cosine function. Figure 2.13 shows an inverse sine phase modulator, which consists of a balanced modulator, a summer,

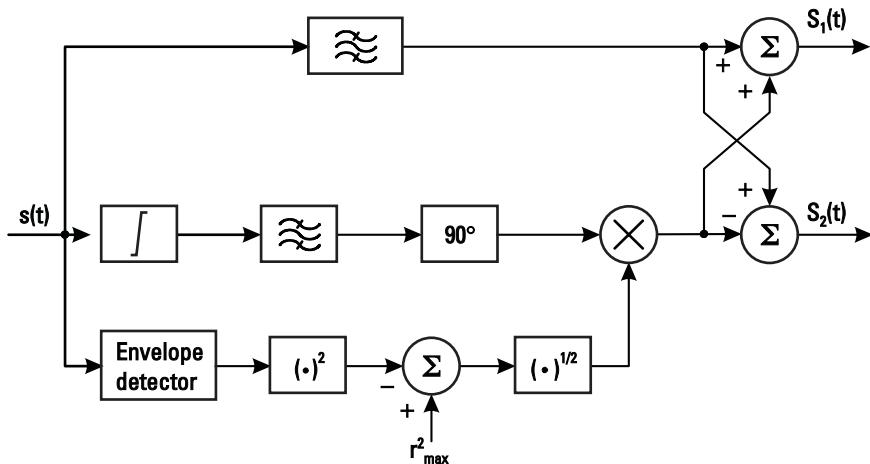


Figure 2.12 Signal component separation with the in-phase/quadrature method [10].

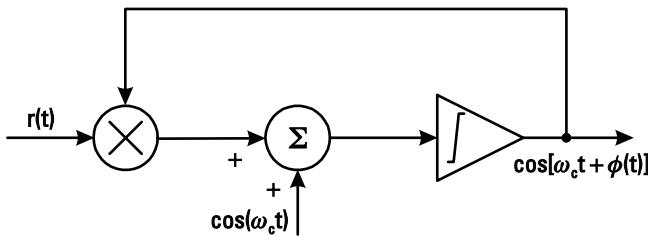


Figure 2.13 Inverse sine phase modulator with feedback loop [18].

and an envelope-limiting amplifier [18]. The circuit takes a baseband envelope as the input and modulates the phase of an IF carrier approximately in accordance with the *arcsin* characteristic. The modulator output and its conjugate can then be used along with the original PM of the input signal to provide the two constant envelope signals. The analysis [18] shows if the loop delay τ is small and satisfies

$$\omega_c \tau = n\pi + \frac{\pi}{2}, \quad (2.29)$$

where ω_c is the carrier radian frequency and n is an arbitrary integer; then the phase of the modulator output consists of a linear term and a distortion term proportional to τ^2 such that

$$\phi(t) \approx (-1)^n \arcsin \left\{ r(t) + \frac{r(t)r'(t)}{2[1 - r^2(t)]} \tau^2 \right\} \quad (2.30)$$

where $r'(t)$ is the derivative of the signal envelope. The second-order effect of the loop delay in the above expression could result in a considerable improvement on distortion and bandwidth. Note that no filter is used in this modulator.

Another analog technique for generation of the quadrature signal is illustrated in Figure 2.14 [19]. The AM and PM of the input are separated and summed together to pass through a comparator or a hard limiter, whose output is either $+1$ or -1 —the polarity of the driving signal. The in-phase component of the quadrature signal is formed after bandpass filtering the comparator output; that is,

$$\begin{aligned} S_1(t) &= \frac{4}{\pi} \cos[\omega t + \theta(t)] \int_0^\infty \frac{\cos[r(t)x]}{x} J_1(x) dx \\ &= \frac{4}{\pi} \cos\{\arcsin[r(t)]\} \cos[\omega t + \theta(t)] \end{aligned} \quad (2.31)$$

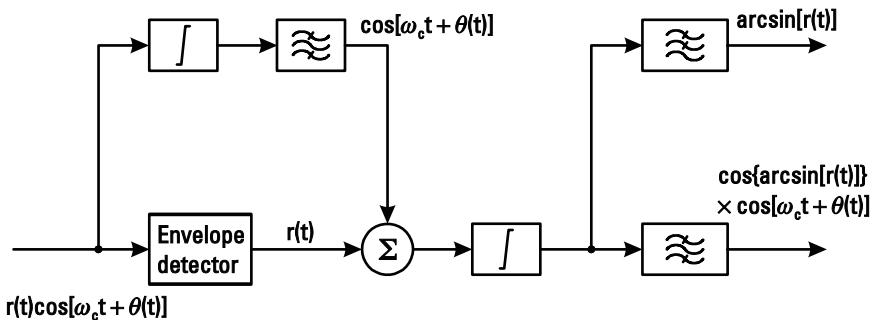


Figure 2.14 Generation of the arcsine and quadrature signal [19].

where J_1 is the first-order Bessel function of the first kind. The envelope $r(t)$ is assumed to be normalized to r_{\max} . As a by-product, the inverse sine is obtained by lowpass filtering the comparator output—that is,

$$\begin{aligned} S_0(t) &= \frac{2}{\pi} \int_0^{\infty} \frac{\sin[r(t)x]}{x} J_0(x) dx \\ &= \arcsin[r(t)] \end{aligned} \quad (2.32)$$

The quadrature signal, as its name implies, always has a 90° phase difference with respect to the input. This fact suggests that the quadrature signal can be generated by rotating the input signal by 90° and with a proper scaling. According to (2.26), the scaling can be realized by a variable gain amplifier (VGA) with a gain factor of

$$\text{gain} = \sqrt{\frac{r_{\max}^2}{r^2(t)} - 1} \quad (2.33)$$

It is readily seen that this scheme demands that the VGA gain approaches infinity as the input envelope approaches to zero. This is of course impossible; hence it is important to ensure that there is no zero-crossing in the modulated signal. Fortunately, this requirement is exactly what is required for some linear modulations, including $\pi/4$ -DQPSK and OQPSK. Figure 2.15 shows an implementation of this VGA-based scheme with power feedback [20]. The power of the two component signals are added and then subtracted from a reference signal. The error signal is fed to a lowpass filter and an amplifier to drive the VGA. In other words, the feedback loop ensures the constant envelope of the two output signals. The SCS chip was

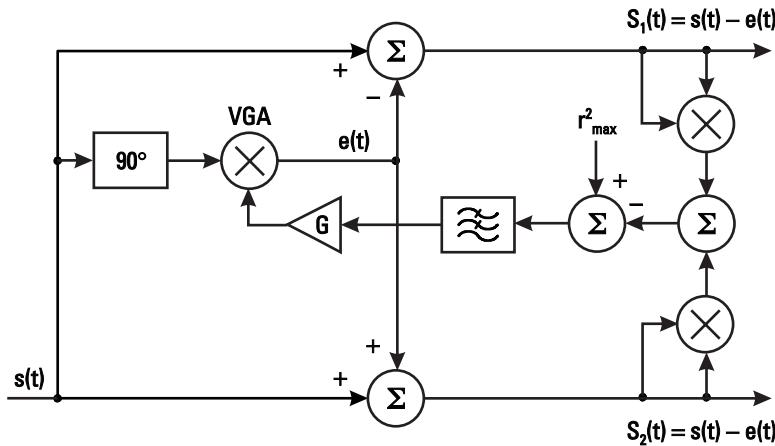


Figure 2.15 SCS with power feedback [20].

implemented in a $0.8\text{-}\mu\text{m}$ BiCMOS technology and operates at 200-MHz IF. The tested bandwidth is 1.8 MHz, due to on-chip implementation of the SCS.

The VGA gain may be directly calculated based on (2.33), and that results in an improved bandwidth over the feedback approach. Figure 2.16 illustrates the block diagram of this scheme [21, 22]. The VGA control signal is obtained by the use of a squaring circuit, a lowpass filter, a reference signal, and a key analog circuit that performs division and “square-root” functions.

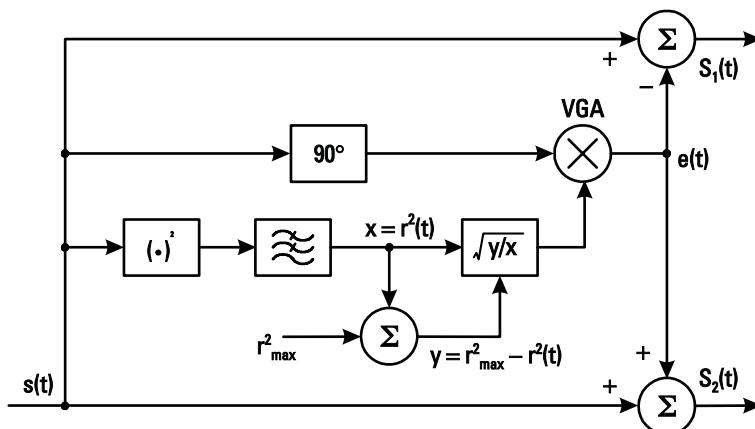


Figure 2.16 SCS without feedback loop [21].

The key circuit is implemented based on a translinear core, as shown in Figure 2.17. Six NPN bipolar transistors Q_1-Q_6 form a translinear loop. The translinear principle states that in a translinear loop the product of the collector current density in the clockwise direction is equal to the product of the collector current density in the counterclockwise direction [23], which can be easily proved by adding the base-emitter voltages of the transistors in each direction. The diode-connected transistors like Q_1 and Q_6 are used to take the square of the current I_c and I_{out} , which transforms the control current to a voltage and takes the square root, respectively. Thus

$$I_{\text{out}} = \sqrt{\frac{I_c^2(I_{\text{ref}} - I_{\text{sq}})}{I_{\text{sq}}}} \quad (2.34)$$

A VGA control voltage proportional to the output current is then generated. The major error source of the above expression is the finite current gain of the transistors, particularly as I_{sq} —proportional to the instantaneous input level—approaches a small value. Some compensation circuits can be used to improve the accuracy of the calculation [21]. The SCS chip set of [21] was

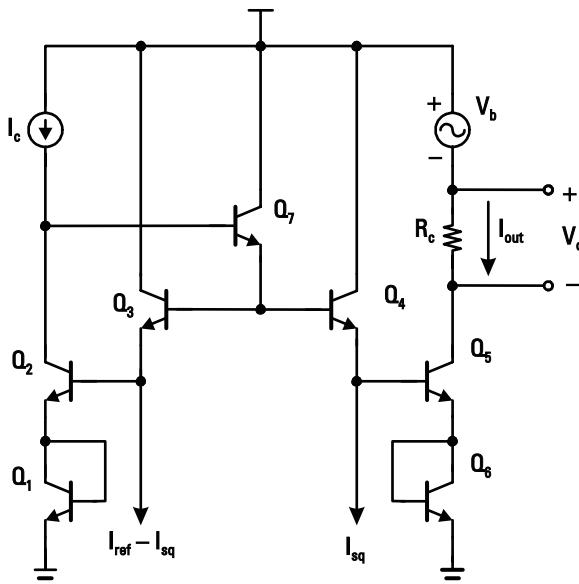


Figure 2.17 A bipolar translinear core [21].

fabricated in a 0.8- μm BiCMOS process and operated at a 200-MHz IF. A CMOS SCS chip set of the similar scheme is given by [22], in which the key circuit is based on the voltage translinear circuit principle to perform the square root and division [24].

So far, the discussion of the SCS is restricted to that of an analog circuit implementation. An apparent disadvantage of the analog implementation is the difficulty to obtain highly accurate modulated component signals to meet the stringent linearity requirements imposed by the wireless standards. The advent of modern DSP and the application-specific DSP (ASDSP) technology makes a digital implementation of the SCS feasible [25]. With this method, all the signal processing is implemented at the baseband, and two quadrature modulators are used to translate the two component signals to the desired carrier frequency, as shown in Figure 2.18. However, the computational load required by the DSP imposes a practical constraint on the performance in terms of bandwidth and power consumption. There are also other constraints and trade-offs associated with the DSP/ASDSP approach, such as the word length of the DSP, the sampling rate, and the reconstruction filter selection. These factors will be discussed in Chapter 3. One disadvantage of this approach is that the reconstruction filter cannot be the pulse-shaping filter due to the nonlinear transformation in the SCS.

A digital SCS method based on a look-up-table (LUT) was suggested in [26], in which a two-dimensional LUT stores the complex quadrature signal, and each LUT entry is addressed by the in-phase and quadrature components of the baseband digital input. The only instructions required are a table lookup and two complex additions for every processed sample. A practical

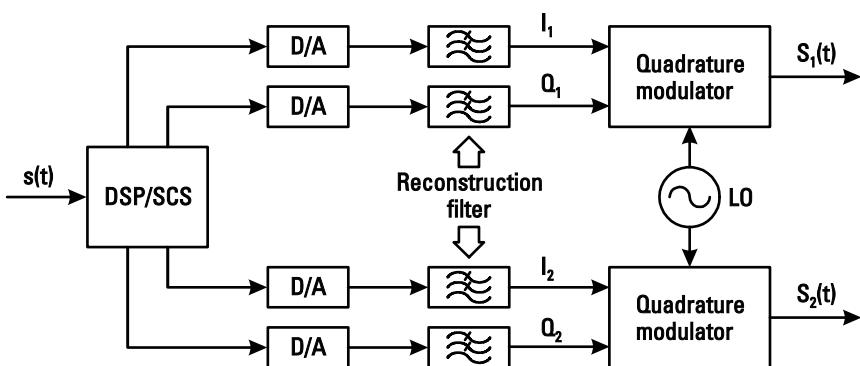


Figure 2.18 Digital signal component separator with frequency translation.

constraint of this method is the extremely large memory requirement on the LUT. As an example, if the complex input and quadrature signals are quantized to nine bits, the LUT would have at least $2^9 \times 2^9 \times 2 \times 9 \approx 4.7$ Mbits memory to store all the possible values of the quadrature signal. However, by utilizing the symmetry along the real and imaginary axes, the memory requirement would be reduced by a factor of four. Note that a further advantage of symmetry can be taken, as shown in Figure 2.19. The vector s' and s are symmetric with respect to the 45° axis, and their quadrature vector e' and e are symmetric with respect to the -45° axis, according to the following expressions,

$$\theta + \theta' = 90^\circ \quad (2.35)$$

$$\alpha + \alpha' = \theta + \theta' + 270^\circ \quad (2.36)$$

Hence, the following relations can be readily derived

$$s'_I = s_Q \quad (2.37)$$

$$s'_Q = s_I \quad (2.38)$$

$$e'_I = -e_Q \quad (2.39)$$

$$e'_Q = -e_I \quad (2.40)$$

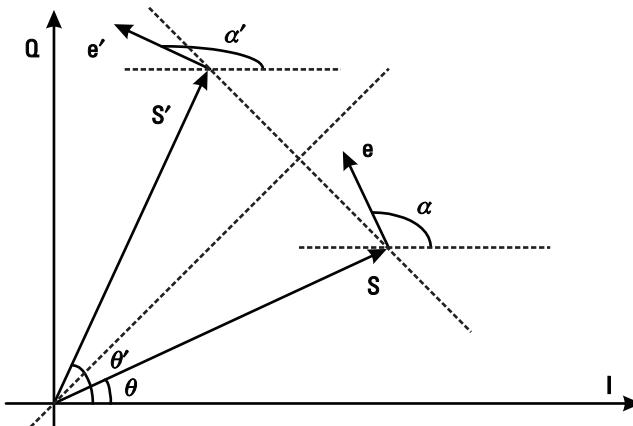


Figure 2.19 LUT symmetry with respect to the 45° axis.

which further reduces the total memory requirements. The total memory reduction of a factor of eight would result in approximately 590 Kbits—still a very large number, but not prohibitive for many applications.

An improved approach, using a one-dimensional LUT, was proposed in [27]. The LUT stores the amplitude-scaling factor $\sqrt{r_{\max}^2 / r^2(t) - 1}$ —the equivalent VGA gain in (2.33) and the table entry is addressed by the input power $r^2(t)$. The real and imaginary parts of the quadrature signal $e(t)$ are obtained by rotating the input vector $s(t)$ by 90° and multiplying it by the addressed scaling factor. Simulation can be used to minimize the table resolution (the number of table entries) and precision (bit length of the table entry) for the required linearity. As an example, less than 14 Kbits of memory was required to achieve -70 -dBc ACI with two Class F power amplifiers. The huge reduction in the required memory is achieved at the expense of an increased algorithmic overhead, for now several multiplication instructions are required for each processed sample.

2.5 Path Imbalance and Its Effects on Linearity

The path imbalance is a manifestation of the gain and phase mismatch between the two parallel power amplifier paths. This imbalance mainly comes from the mismatch between the gain and phase characteristics of the two power amplifiers, possibly arising from the impedance-matching and biasing circuits as well as mismatches in the individual transistors themselves. The nonideality of the power combiner and the misalignment of the metal wires and transmission lines also contributes to the path imbalance. When the matching components are realized on an integrated circuit with lithographic techniques, a variety of effects can cause the effective size of the components to be different from the size of the layout mask. A matching error of 0.1 to 1% can be achieved with careful layout. Matching accuracy becomes worse with discrete components, and at best 3 to 5% matching can be achieved. Due to the constant envelope feature of the power amplifier driving signals, the path imbalance is straightforward to analyze in terms of the amplitude and phase difference between the two amplified component signals. The path imbalance has been investigated in [28, 29] and needs to be corrected for most practical applications, since it contributes directly to ACPR and EVM degradation. This section derives a closed-form expression for the out-of-band spectral regrowth for a two-tone test. A simple means to estimate the ACI for a linearly modulated signal is also given.

2.5.1 Two-Tone Linearity Analysis of an Outphased Amplifier with Path Mismatch Effects

The classic two-tone test is the simplest and yet most effective method to evaluate the nonlinearity of an amplifier and can illustrate both magnitude and phase distortion. The two-tone test is almost universally accepted, although alternative techniques such as white noise or multicarrier test signals are more suitable for characterizing the nonlinearities in digital modulation transmitters. During the two-tone test, two equal-amplitude sinusoidal signals are applied to the amplifier, and the signal envelope experiences complete variation between zero and the maximum value, by means of which the amplifier characteristic is examined. Because of the zero-crossing envelope, the ACI performance excited by two-tones is generally worse than the ACI generated by some linearly modulated signal with modest envelope variation. If the two signal frequencies are assumed to be at frequencies $\omega_0 \pm \Delta\omega$, we have the complex input signal

$$s(t) = Ae^{j(\omega_0 - \Delta\omega)t} + Ae^{j(\omega_0 + \Delta\omega)t} \quad (2.41)$$

where A is the amplitude of each tone. The instantaneous envelope of the input two tone is given by

$$\begin{aligned} r(t) &= \sqrt{s(t)s^*(t)} \\ &= 2A|\cos \Delta\omega t| \end{aligned} \quad (2.42)$$

The two-tone signal level should be arranged such that the peak envelope power (PEP) is equal to that of the full-power rating of the amplifier, (*i.e.*, $r_{\max} = 2A$ in this case). Hence the ideal quadrature signal $e(t)$ is calculated by

$$\begin{aligned} e(t) &= js(t)\sqrt{\frac{r_{\max}^2}{r^2(t)} - 1} \\ &= js(t)|\tan \Delta\omega t| \end{aligned} \quad (2.43)$$

Now, we introduce the effects of path mismatch on the amplifier performance. We assume that the gain and phase delay of the power amplifier branches are G_0 and ϕ_0 and that the gain and phase imbalance of the bottom amplifier branch with respect to the upper one are $\Delta G/G_0 \ll 1$ and $\Delta\phi \ll 1$, respectively. The final output signal would be the desired signal plus an interference term

$$\begin{aligned} S(t) &= G_0 e^{j\phi_0} [s(t) - e(t)] + G_0 \left(1 + \frac{\Delta G}{G_0}\right) e^{j(\phi_0 + \Delta\phi)} [s(t) + e(t)] \\ &\approx 2G_0 e^{j\phi_0} s(t) + G_0 e^{j\phi_0} \left(\frac{\Delta G}{G_0} + j\Delta\phi\right) e(t) \end{aligned} \quad (2.44)$$

The quadrature signal $e(t)$ is a wideband signal, and its spectrum in the two-tone case can be easily calculated. Rearranging $e(t)$, we have

$$\begin{aligned} e(t) &= 2A \cos \Delta\omega t |\tan \Delta\omega t| j e^{j\omega_0 t} \\ &= r_{\max} e_s(t) j e^{j\omega_0 t} \end{aligned} \quad (2.45)$$

where

$$e_s(t) = \cos \Delta\omega t |\tan \Delta\omega t| \quad (2.46)$$

Hence $e(t)$ can be regarded as a sinusoidal signal modulated by a slow variation term $e_s(t)$. Taking the Fourier series expansion of $e_s(t)$, we have

$$e_s(t) = \sum_{n=-\infty}^{\infty} a_n e^{jn\Delta\omega t} \quad (2.47)$$

where the Fourier series coefficient is given by

$$\begin{aligned} a_n &= \frac{1}{T} \int_T e_s(t) e^{-jn\Delta\omega t} dt \\ &= \frac{2 \sin(n\pi/2)[n - \sin(n\pi/2)]}{\pi n^2 - 1} \end{aligned} \quad (2.48)$$

Obviously all even-order terms remain zero, and only the odd-order terms are left. Both out-of-band spurs and inband distortion are generated—the $n = \pm 1$ -order terms are inband distortion, and the rest of the $n = \text{odd}$ -order terms are out-of-band interference. A close examination reveals that the periodic 180° phase discontinuity of the two-tone signals as the envelope reaches zero is responsible for the out-of-band spurs. Clearly, the most significant spurs occur at $\omega_0 \pm 3\Delta\omega$, in analogy to intermodulation distortion in a weakly nonlinear amplifier.

In a two-tone test, the ACI, or out-of-band rejection, is the ratio of the magnitude of the most significant spurs to the magnitude of the desired signals. In this case, ACI is given by

$$\begin{aligned}
 \text{ACI} &= 20 \log_{10} \left[\frac{1}{\pi} \sqrt{\left(\frac{\Delta G}{G_0} \right)^2 + \Delta\phi^2} \right] \\
 &= 10 \log_{10} \left[\left(\frac{\Delta G}{G_0} \right)^2 + \Delta\phi^2 \right] - 9.9 \text{ (dB)} \quad (2.49)
 \end{aligned}$$

Figure 2.20 displays the spectrum of the outphasing system output, in the presence of 2° phase imbalance and 0.3-dB gain imbalance. It is interesting to note that the adjacent two spurs have the same magnitude and opposite sign.

2.5.2 ACI Estimation with Gain and Phase Mismatch

In an outphasing system, the band-limited input is decomposed into two component signals, then amplified and recombined. In the presence of gain and phase imbalance, the combined output would be

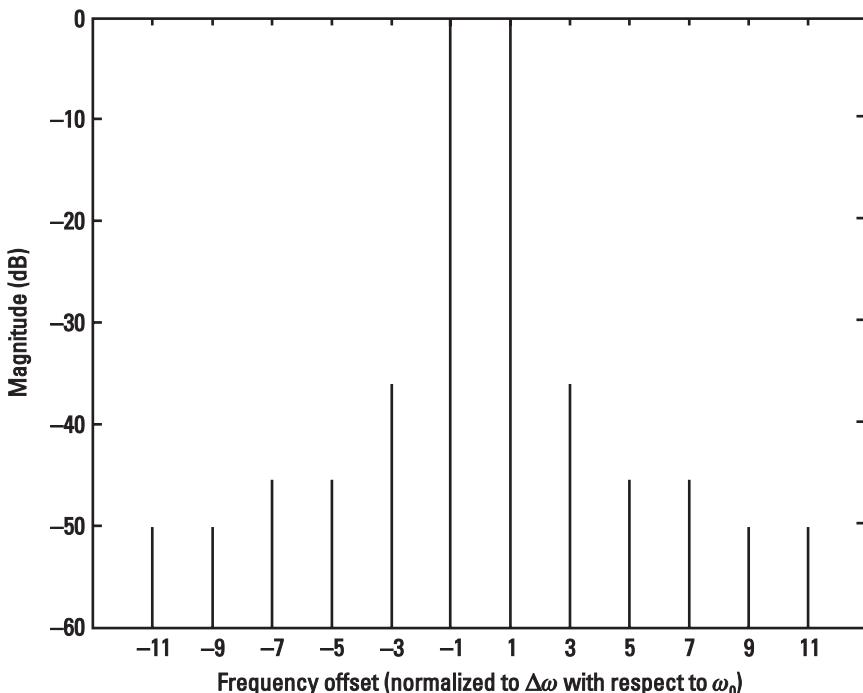


Figure 2.20 Output spectrum of a two-tone test of outphasing system with 2° phase imbalance and a 0.3-dB gain imbalance.

$$\begin{aligned}
S(t) &= S_1(t) + S_2(t) \\
&= [s(t) - e(t)] + \left(1 + \frac{\Delta G}{G_0}\right) e^{j\Delta\phi} [s(t) + e(t)] \\
&= \underbrace{\left[1 + \left(1 + \frac{\Delta G}{G_0}\right) e^{j\Delta\phi}\right] s(t)}_{\text{desired term}} + \underbrace{\left[\left(1 + \frac{\Delta G}{G_0}\right) e^{j\Delta\phi} - 1\right] e(t)}_{\text{error term}} \quad (2.50)
\end{aligned}$$

The signal has been normalized to $G_0 e^{j\phi_0}$ —the gain and phase characteristics of the upper amplifier branch. From (2.50), it is clear that the gain and phase imbalance modify the magnitude and phase of the output signal. The first term in (2.50) constitutes the desired output signal with slightly modified amplitude and phase, while the second term creates both the out-of-band interference and inband distortion. Equation (2.50) can be simplified and rearranged in the case of a small amount of the gain and phase imbalance.

$$S(t) \approx 2s(t) + \left(\frac{\Delta G}{G_0} + j\Delta\phi\right) e(t) \quad (2.51)$$

The power spectrum is thus the weighted sum of the desired spectrum of $s(t)$ and the spectrum of $e(t)$. The spectrum of the quadrature signal $e(t)$ is difficult to calculate and strongly dependent on the modulation—knowing the bandwidth of the source signal is not adequate to predict the bandwidth of the quadrature signal. The quadrature signal is generally wideband, and its spectrum extends far into adjacent and alternate channels [30]. The second term in (2.51) is evidence of the incomplete cancellation of the quadrature signal during power combining, and it is responsible for the out-of-band spectrum regrowth. Note that, in the case of small amount of gain and phase imbalance, the inband distortion generated by this term hardly affects the spectrum characteristics of the desired signal. Thus, a simple way to estimate the out-of-band rejection is to do a simulation of the quadrature signal and compute its spectrum first, then compare it to the spectrum of the source signal and obtain the ratio of the peak spectral density outside the channel of the quadrature signal spectrum to the inband spectral density of the source signal—denoted as P_e . Finally, the ACI of the output signal can be approximated by

$$\begin{aligned}
 \text{ACI} &= 20 \log_{10} \left[\frac{1}{2} \sqrt{\left(\frac{\Delta G}{G_0} \right)^2 + \Delta\phi^2} \times P_e \right] \\
 &= 10 \log_{10} \left[\left(\frac{\Delta G}{G_0} \right)^2 + \Delta\phi^2 \right] + P_e - 6.0 \quad (\text{dB}) \quad (2.52)
 \end{aligned}$$

Figure 2.21 displays the spectrum of the desired signal as well as the quadrature signal for various modulations, filtered with the square-root raised cosine with a roll-off factor of 0.35. The sampling rate was 16× higher than the symbol rate. The inband power spectrum density has been normalized to 0 dB, and the spectra of the quadrature signals are offset accordingly. It can be seen that for the quadrature signal, a significant portion of the power extends into adjacent and alternate channels. One aspect of this is that an input signal with a wider range of envelope variations—such as 16-QAM—results in a lower PAP ratio on the quadrature signal. The result of this is a higher inband distortion power for the quadrature signal. As we can see, P_e is roughly -9 dB for QPSK, OQPSK and $\pi/4$ -DQPSK, and -6 dB for 16-QAM. With this value in mind, it is easy to estimate the ACI of

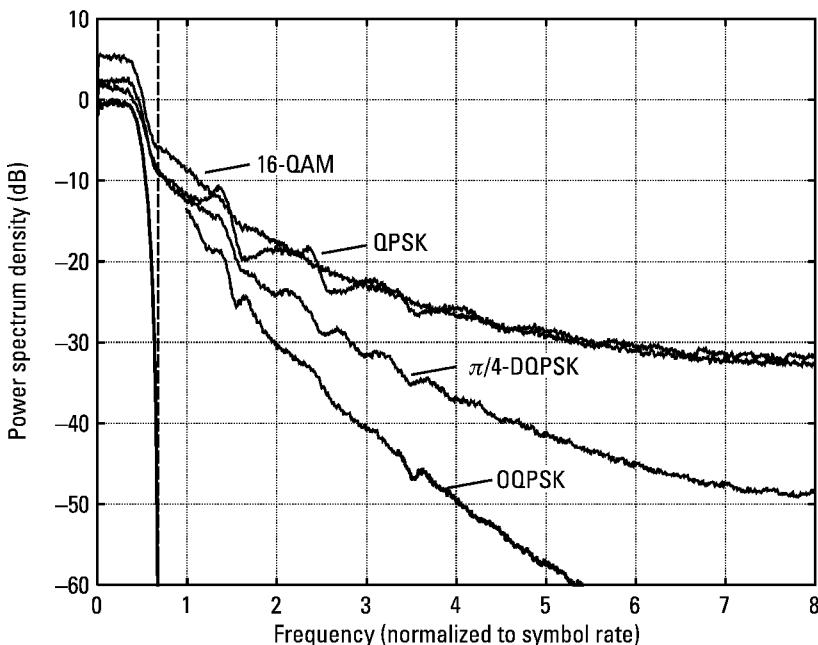


Figure 2.21 Simulated spectrum for $s(t)$ and $e(t)$ for various modulations.

the outphasing system with the knowledge of the path imbalance. Compared to the two-tone case, we see that generally the ACI for bandlimited signals is 2–5 dB better than that for the two-tone case, depending on the modulation.

Figure 2.22 shows similar spectra for the CDMA IS-95 uplink. The modulation here is OQPSK. Two spectra with different matching conditions are also illustrated for comparison. One spectrum is obtained under a 10% gain imbalance and a 10° phase imbalance, and the other is obtained under a 3.5% gain imbalance and a 2° phase imbalance. Their spectra are normalized to have the same peak inband density. According to (2.52) the ACI with mismatching is -25 dBc and -38 dBc, respectively, which is in good agreement with the simulation results.

2.6 Effect of Quadrature Modulator Errors on Linearity

Modern mobile systems frequently use a quadrature modulator to generate complex vector modulation. To translate the complex baseband waveform to

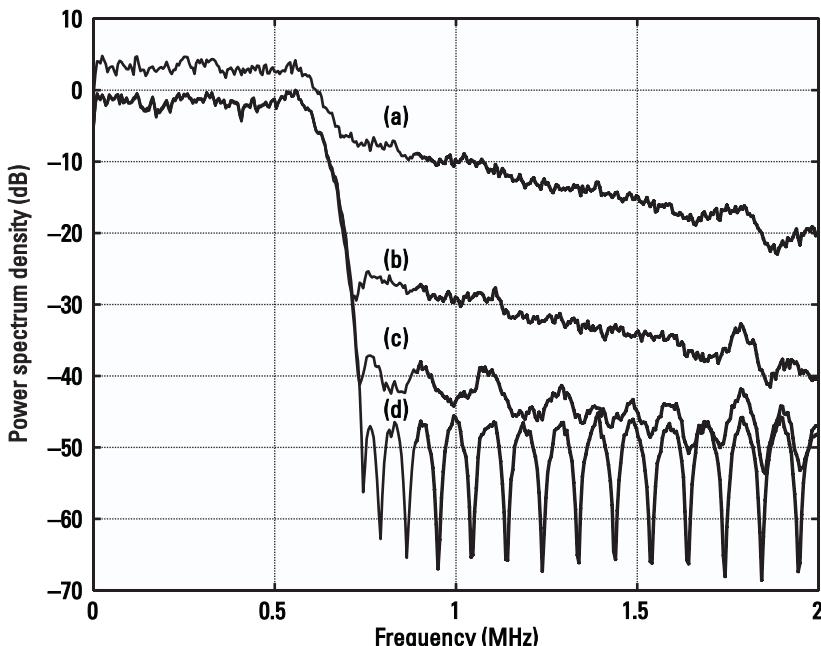


Figure 2.22 Simulated spectrum for CDMA IS-95: (a) quadrature signal, (b) total output with 10% gain imbalance and 10° phase imbalance, (c) total output with 3.5% gain imbalance and 2° phase imbalance, and (d) ideal input signal.

the desired carrier frequency band, two general topologies are used in a transmitter—indirect and direct upconversion. An indirect upconversion system implements the modulation at an IF and then translates that result to a higher frequency at the RF, requiring a local oscillator (LO) and IF filtering. A direct upconversion system performs the data modulation and frequency translation at the same time, hence eliminating the IF block. The resultant more compact structure and reduced power consumption is suitable for monolithic integration, but care must be taken to keep the resulting spurious signals outside the band of interest.

The block diagram of a typical quadrature modulator is shown in Figure 2.23, in which the balanced baseband I/Q waveforms modulate two carriers driven in quadrature phase, and the resultant signals are summed. Various factors and implementation imperfections contribute to the degradation of the resulting modulation accuracy, and the effects can be collectively represented by the gain error, phase error, and LO leakage [31]. The quadrature modulator gain error describes the gain mismatch between the I and Q channels. The phase error mainly comes from the imperfection of the 90° phase shifter. The LO-RF feedthrough causes LO leakage. Besides, the DC offsets in the I/Q channels give rise to a similar effect to LO leakage, and hence are ascribed to that term. One way of visualizing the impairment of quadrature error is to consider a baseband sinusoidal single-tone input. The modulator output trajectory in the complex I/Q plane becomes an ellipse—rather than an ideal circle—in the case of gain and phase error. The phase error further rotates the principal axes of the ellipse by 45° . The LO leakage term simply shifts the whole trajectory by an equal amount. By use of the complex envelope notation in [31], the complex envelope of the modulated RF signal can be derived as follows

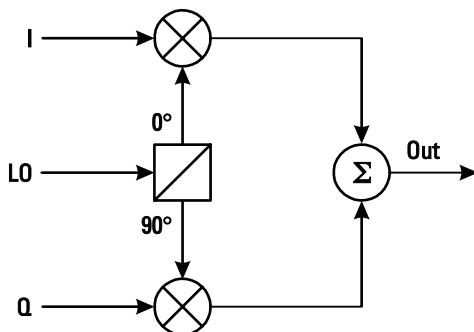


Figure 2.23 Quadrature modulator block diagram.

$$R(t) = \frac{1}{2}(1 + C)A(t) + \frac{1}{2}(1 - C)A^*(t) + p \quad (2.53)$$

where $A(t)$ is the complex baseband waveform and $A^*(t)$ is the complex conjugate. The constant p is a complex number to represent the LO leakage term. The constant C characterizes the gain and phase mismatch of the quadrature modulator in the Q channel with respect to the I channel; that is,

$$C = (1 + g)e^{j\delta} \quad (2.54)$$

where g is the gain error and δ is the phase error. In the case of the small quadrature error, (2.53) approximates to

$$R(t) = \left[1 + \frac{1}{2}(g + j\delta) \right] A(t) - \frac{1}{2}(g + j\delta)A^*(t) + p \quad (2.55)$$

The first term is the desired signal with a small amplitude and phase modification. The second term is the spurious/image signal, which represents the counterclockwise rotating version of the input. The image rejection ratio is thus given by [31]

$$\text{IR} \approx 10 \log_{10}(g^2 + \delta^2) - 6.0 \quad (\text{dB}) \quad (2.56)$$

2.6.1 Quadrature Modulator Error Minimization

The quadrature modulator error generates a small variation in the amplitude and phase of the modulated signal. When this signal passes through a nonlinear power amplifier, intermodulation terms are generated, and their spectra expand into adjacent channels, causing interference to other users [32]. This effect is readily seen by taking the cube in (2.53) and collecting the A^3 term and $(A^2 A^*)$ term. The leakage power is proportional to the square of the quadrature errors. The improvement of quadrature error is reduced when a digital predistorter is included preceding the quadrature modulator. Even a relatively small amount of quadrature error compromises the improvement gained through predistortion [32, 33]. Quadrature modulator errors have a similar detrimental effect to outphasing power amplifiers; the details will be covered in Section 2.6.2.

Great efforts have been made to reduce the quadrature modulator errors in different aspects. One important research activity applies to

the generation of quadrature LO signals. An RC-CR network is the simplest quadrature signal generator, as shown in Figure 2.24. It is easily proven that the two output signals have 90° phase difference at all frequencies but have equal amplitude only at the pole frequency $\omega = 1/RC$. A polyphase filter, shown in Figure 2.25, is often used as a broadband 90° phase shifter [34]. The number of filter stages is determined by the required bandwidth. This structure is less sensitive to the variations of the component values compared to the simple RC-CR network. One disadvantage is the increased thermal noise. An alternative approach is RC-CR pole frequency tuning, in which varactor diode capacitors or transistor junction capacitors are changed by an appropriate control voltage. Havens' quadrature circuit exploits the fact that the two diagonals of a rhombus are perpendicular to each other [35]. In other words, the summation and difference of two equal-amplitude vectors always have a 90° phase difference, as illustrated in Figure 2.26. The limiters ensure that the two input and resulting signals have equal amplitude; otherwise a phase error will result. It is easy to show that the amplitude mismatch in percentage generates approximately the same amount of phase error in radians. For example, every 0.1-dB amplitude error contributes to a 0.7° phase error. The quadrature oscillator is another way to generate quadrature LO signals with equal amplitude [36].

Despite those efforts, however, the quadrature modulator error is expected to change with environmental variation, channel frequency, temperature, component aging, supply voltage, and biasing. Hence, some sort of continuous tracking and compensation may be necessary to achieve exceptional linearity performance. Several schemes have been proposed to systematically track and correct the quadrature modulator error, including the gain and phase mismatch and carrier leakage [37–40]. These methods use an envelope detector, which checks the modulation output, as a feedback means

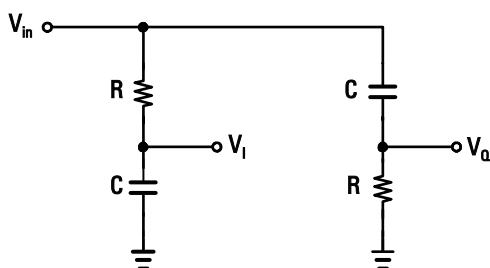


Figure 2.24 Quadrature signal generation with RC-CR network.

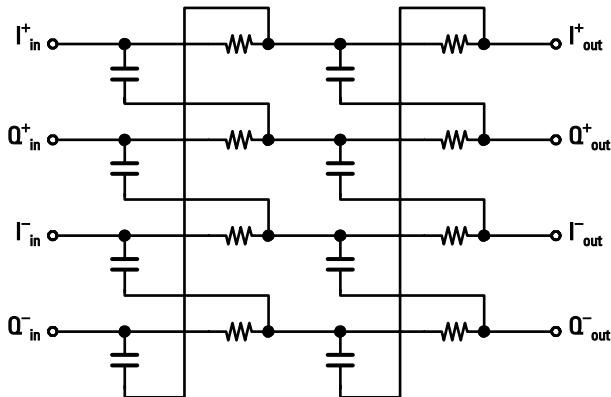


Figure 2.25 A two-stage polyphase filter [34].

to guide the adaption of the digital correction circuit. With the knowledge of quadrature error, the compensation is accomplished by simply adjusting the amplitude, phase, and dc offset of the baseband I and Q signals. The estimation of the quadrature error is based either on the predefined test vectors [37, 39, 40] or on the comparison of the modulation data and the envelope feedback [38, 40], as shown in Figure 2.27(a, b), respectively. The requirement of training in the former methods restrains their application, although they provide relatively simple adaption. The latter methods enable background operation, by use of Newton-Raphson algorithm or least squares algorithm. In addition to their computational load, the algorithm complexity grows greater when quantization error and loop delay have to be taken into account.

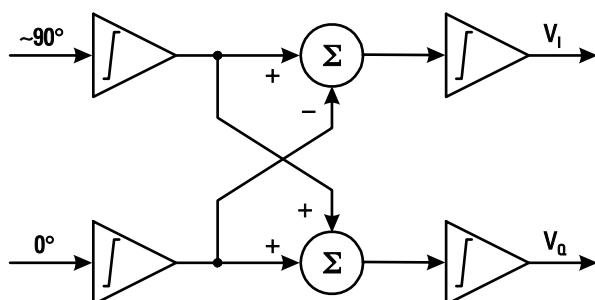


Figure 2.26 Havens' quadrature signal generator [35].

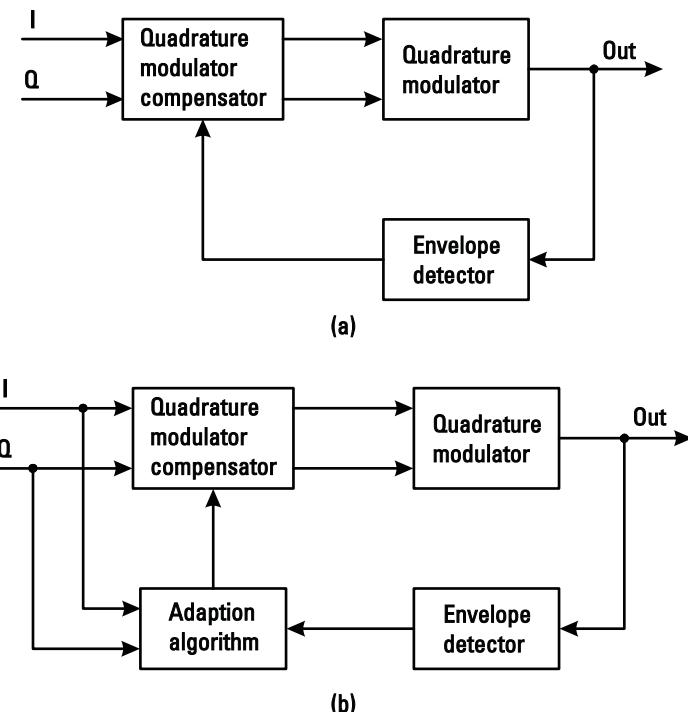


Figure 2.27 (a) Quadrature modulator error correction by test vectors, and (b) with background adaption.

2.6.2 Quadrature Modulator Error Effects on Outphasing Systems

The effect of quadrature modulator error on outphasing power amplifiers is intrinsically different from that of the same error on digital predistorters. In the case of a digital predistorter, the quadrature modulator generates a slightly scaled version of the desired signal and a small amount of counter-rotating signal. Before the modulated signal passes through the nonlinear amplifier, the counter-rotating signal has no effect on the spectral leakage to adjacent channels, and the spectrum is not expanded although the EVM is increased. The excess amount of spectral regrowth results from the quadrature modulator error, in conjunction with the amplifier nonlinearity. By contrast, in outphasing power amplifiers, the band-limited waveform is spectrally expanded *prior* to the quadrature modulator, and the out-of-band rejection relies on the precise cancellation of two equal errors. Consider a simple outphasing system with the power amplifier removed. By taking

the quadrature modulator gain and phase mismatch into account, the output signal can be described by

$$\begin{aligned} S_o(t) &= \frac{1}{4}(1 + C_1)S_1(t) + \frac{1}{4}(1 - C_1)S_1^*(t) \\ &\quad + \frac{1}{4}(1 + C_2)S_2(t) + \frac{1}{4}(1 - C_2)S_2^*(t) \\ &= s(t) + \frac{1}{4}(\alpha_1 + \alpha_2)(s - s^*) - \frac{1}{4}(\alpha_1 - \alpha_2)(e - e^*) \end{aligned} \quad (2.57)$$

where

$$\begin{aligned} \alpha_{1,2} &= C_{1,2} - 1 \\ &\approx g_{1,2} + j\delta_{1,2} \end{aligned} \quad (2.58)$$

The second term— $(\alpha_1 + \alpha_2)(s - s^*)/4$ —will not contribute to the out-of-band spectrum because it is linearly proportional to s . However, since the quadrature modulator errors generally cannot simply cancel out, the third term in (2.57) is responsible for the spectral regrowth even without the presence of a nonlinear power amplifier. The leakage power is proportional to the square of the quadrature modulator error. This simple analysis shows that quadrature modulator error probably has more detrimental effects on outphasing power amplifier performance than the same error has on other linearization techniques.

A quadrature modulator error degrades the constant envelope characteristics of the two component signals in an outphasing system. While the power amplifier may be driven into saturation for maximum efficiency, the variation of the signal envelope nevertheless creates a devastating effect on the overall system performance. The linearity degradation is dependent upon both quadrature modulator error and amplifier nonlinearity. The envelope variation can be estimated by examining (2.53). It can be shown that in the case of small quadrature gain and phase mismatch, the envelope of the first two terms in (2.53) is bounded by

$$\begin{aligned} \left(1 + \frac{1}{2}g - \frac{1}{2}\sqrt{g^2 + \delta^2}\right)|A| &\leq |(1 + C)A + (1 - C)A^*| \\ &\leq \left(1 + \frac{1}{2}g + \frac{1}{2}\sqrt{g^2 + \delta^2}\right)|A| \end{aligned} \quad (2.59)$$

Hence, the worst-case envelope variation in the presence of quadrature gain and phase mismatch is given by

$$\left| \frac{\Delta R}{R} \right| \leq \frac{1}{2} \sqrt{g^2 + \delta^2} + |\rho| \quad (2.60)$$

As an example using (2.60), with 0.3-dB gain error, 2° phase error, and -30-dBc LO leakage, we get an envelope variation of roughly $\pm 6\%$ or ± 0.5 dB. Such an envelope fluctuation may or may not cause the power amplifier to generate a significant amount of spectral regrowth, depending on the selection of the power amplifier. If the amplifier generates little spectral regrowth due to input envelope variation, then the overall spectral regrowth of the system may be estimated with (2.57).

A semianalytical treatment is given in [41], which elaborates on the coupled effects of the quadrature modulator error, path imbalance, and the power amplifier nonlinearity. This method originates from [32] in the case of constant envelope modulation such as GMSK. The procedure expresses the amplifier nonlinearity into a memoryless Taylor series, substitutes the modulated signal for each amplifier branch, sums the branches together, and finally identifies the dominant distortion terms. The contribution of each term is calculated and summed together. Since the power amplifier operates at rather modest input dynamic range— ± 0.5 dB in the previous example—it is straightforward and sufficient to represent it with a power series by the measured AM-AM and AM-PM characteristics. A power series expansion to the fifth order is usually sufficient. Most digital modulations are based on some sort of symmetrical quadrature constellation; as a result many cross-correlation products become zero. Nonzero crosscorrelation terms are typically not dominant, unless they are almost the same size and no other terms dominate. By ruling out almost all the cross terms, the spectrum is simply the power addition of each weighted term, and the result is usually close to the real simulation within 1 dB. The paper shows a strong detrimental effect of quadrature modulator error—with a 0.1-dB gain error, 0.3° phase error, and a -40-dBc carrier leakage, which represents the state-of-the-art technology available today; the resultant spectral regrowth with a Class C power amplification is around -50 dBc, still rather high.

Two points should be addressed. First, in contrast to the band-limited baseband inputs applied to a quadrature modulator in the usual case, the quadrature modulator in an outphasing system is producing a constant envelope output signal. This may result in a simpler approach for tracking and correcting quadrature modulator errors than is applicable in the more

general case. Second, since the information is carried on the zero-crossing of two component signals, two limiters may be used to flatten the envelope fluctuation due to quadrature modulator error, before the signals are sent to the power amplifiers. The limiter used must have sufficient bandwidth so as not to introduce extra AM-PM distortion. The limiter approach may result in improved spectral purity with little added expense. The use of switching-mode power amplifiers such as Class D or Class E in replacement of the saturated Class B or Class C amplifier may benefit from a similar rationale.

2.7 SCS Quantization Error Effects on Outphasing Systems

The rapid evolution of modern DSP technology makes it an attractive candidate for the implementation of the SCS in an outphasing power amplifier system. With this approach, the signal separation is done digitally in baseband and converted to analog waveforms prior to upconversion and power amplification. Because of the finite word-length representation of the data stream in the quantizer, quantization noise is inevitably generated. It is important to understand this mechanism and how the quantization noise affects the linearity performance of the outphasing system. By use of application specific DSPs (ASDSPs), the algorithm of the arithmetic blocks and registers can be minimized, and the word length can be optimized to reduce the hardware requirements. This will result in reduced power consumption and cost, and increased bandwidth.

During signal component separation, the quadrature signal is being added to, and subtracted from, the source signal to obtain two constant-envelope component signals. A direct consequence of a DSP implementation of this process is that both the source signal and the quadrature signal are quantized. The source signal is quantized prior to the SCS; thus the amplifier output is a scaled replica of the quantized source signal and the quantization noise sets the final output noise floor. The quantization of the quadrature signal has a different mechanism that gives rise to a random envelope ripple on the power amplifier driving signal. This effect is further enhanced by the nonlinear power amplifier through the AM-AM and AM-PM conversions, which are not canceled. Consequently, the overall SNR is further degraded.

2.7.1 Error Effects of Quantization of the Source Signal

The quantization error of a complicated signal is generally considered to be a stationary random process and is uncorrelated with the original signal [42].

The probability density is uniformly distributed over the range of quantization error. Consequently, the quantizer can be modeled as an additive white noise source. The SNR after a B_s -bit complex quantizer is given by [43]

$$\text{SNR}_s = 6.02B_s + 1.76 - \text{PAP} \quad (\text{dB}) \quad (2.61)$$

where the PAP can be calculated from the probability density function,

$$\text{PAP} = -10 \log_{10} \left[\int_0^1 \rho(r) r^2 dr \right] \quad (\text{dB}) \quad (2.62)$$

where ρ is the probability density function of the modulation and r is the normalized signal amplitude. Note that the normalized average signal power is equal to the inverse of the PAP of the signal. The noise power in (2.61) is doubled in a complex quantizer, due to the fact that the quantization takes place independently on the real and imaginary parts of the source signal.

Considering that the signal power is distributed within the bandwidth of B , which is the single-sided equivalent noise bandwidth of the modulation, while the additive white noise power is distributed up to the bandwidth of half sampling rate, the following expression is readily proven [43]

$$\text{ACI}_s = \frac{2B}{f_s \cdot \text{SNR}_s} \quad (2.63)$$

where f_s is the sampling rate. The equivalent noise bandwidth is dependent on the modulation and baseband pulse shaping. As shown previously, the noise bandwidth of a modulated signal with square-root raised cosine filtering is equal to half of the symbol rate. According to (2.63), for every one-bit increase in word length, the SNR and hence ACI will be improved by 6 dB, while doubling the sampling frequency improves ACI by 3 dB.

2.7.2 Error Effects of Quantization of the Quadrature Signal

When the quadrature signal $e(t)$ is quantized, the quantization error Δe seems to cancel out during the signal combining, as illustrated in Figure 2.28. However, since nonlinear power amplifiers are used instead to improve the power efficiency, the quantization error results in a variation of the magnitude of the signal applied to the power amplifiers, and consequently, the power amplifier gain and phase characteristics are changed due to AM-AM

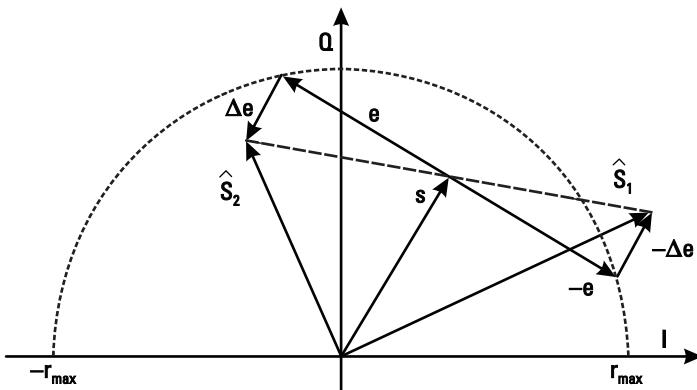


Figure 2.28 Effects of an quantization error Δe on quadrature signal $e(t)$.

and AM-PM conversion. Assuming the changes of the gain of the two amplifiers are ΔG_1 and ΔG_2 , respectively, the amplified output signals are

$$\hat{S}_1 = (s - e - \Delta e)(G + \Delta G_1) \quad (2.64)$$

$$\hat{S}_2 = (s + e + \Delta e)(G + \Delta G_2) \quad (2.65)$$

where G is the gain of the amplifiers evaluated at the constant envelope r_{\max} without quantization. The gain G is a phasor, including both the AM-AM and AM-PM conversion. The quantization error Δe causes both the amplitude deviation and phase deviation of the component signal from the desired value. The effect of the amplitude ripple is enhanced by the nonlinear power amplifiers. However, the amplitude ripple and the phase ripple directly caused by the quantization error Δe cancel out in the first order. This effect is evident from the combined output without the explicit term of Δe ,

$$\begin{aligned} \hat{S}_{\text{out}} &= \hat{S}_1 + \hat{S}_2 \\ &\approx 2G \cdot s + (\Delta G_1 + \Delta G_2) \cdot s - (\Delta G_1 - \Delta G_2) \cdot e \end{aligned} \quad (2.66)$$

The error term is evaluated by assuming that the quantization error is uniformly distributed. Then the average noise power is given by [43]

$$\sigma_e^2 = \frac{1}{3 \cdot 2^{2B}} \int_0^1 \rho(r) |G'(r)|^2 \{1 - \cos[4 \arccos(r)]\} dr \quad (2.67)$$

and the signal power is

$$\sigma_s^2 = \int_0^1 \rho(r) |2rG(r)|^2 dr \quad (2.68)$$

The SNR due to quadrature signal quantization becomes

$$\text{SNR}_q = \frac{\sigma_s^2}{\sigma_e^2} \quad (2.69)$$

Again assuming that the quantization error appears as white additive noise, the resulting ACI is given by

$$\text{ACI}_q = \frac{2B}{f_s \cdot \text{SNR}_q} \quad (2.70)$$

In a similar manner to the source signal quantization, for every one-bit increase in the word length representation, the SNR and hence ACI will be improved by 6 dB. Doubling the sampling frequency improves ACI by 3 dB.

The quantization error is generally a small fraction of the signal magnitude. For example, the amplitude variation will be ± 0.07 dB after a 7-bit quantizer. With this range of ripple, the power amplifier AM-AM and AM-PM distortion can be well-approximated by a straight line. Assuming that the amplifier has a complex gain— $G_r(r)e^{j\theta(r)}$, we can calculate

$$|G'(r)|^2 = G_r'^2 + \theta'^2 \quad (2.71)$$

The quantization noise power is then approximated by

$$\sigma_e^2 = \frac{1}{3 \cdot 2^{2B}} |G'(r)|^2 \int_0^1 \rho(r) (r^2 - r^4) dr \quad (2.72)$$

and the SNR is

$$\text{SNR}_q = \text{SNR}_s - 10 \log_{10} \left[\left(\frac{G'_r}{G_r} \right)^2 + \theta'^2 \right] - 10 \log_{10} \int_0^1 \rho(r) (r^2 - r^4) dr \quad (\text{dB}) \quad (2.73)$$

The SNR could be infinity in the case of perfectly linear power amplifiers. This result shows that it is important to design power amplifiers with nearly linear AM-AM and nearly flat AM-PM characteristics near the operating

point, such that the quadrature signal quantization will not degrade the overall noise performance. Switching-mode power amplifiers may perform better than the classical saturated power amplifiers in this regard, as long as the bandwidth is sufficiently large. It can be shown that the SNR in this case is given by

$$\text{SNR}_q = \text{SNR}_s - \text{PAP} - 10 \log_{10} \int_0^1 \rho(r)(r^2 - r^4) dr \quad (\text{dB}) \quad (2.74)$$

2.8 Linearity Effects of Reconstruction Filter and DSP Sampling Rate

As shown in Figure 2.18, in a digital SCS the signal separation is accomplished inside the processor, and the resultant digital streams are fed to the digital-to-analog (D/A) converters and reconstruction filters to be transformed into analog waveforms. The sampling rate of the DSP is a crucial design parameter and should be kept as low as possible to accommodate the high modulation bandwidth and lower the power consumption. However, a low sampling rate increases the complexity of the reconstruction filter since a low filter order generally requires a high sampling rate and vice versa. Thus, these conflicting requirements must be properly traded off.

The effect of the D/A conversion in the absence of the quantization error is equivalent to the simple block diagram shown in Figure 2.29, in which the original analog waveform is multiplied by an impulse train, followed by a zero-order hold circuit. The sampling images are generated during the impulse train modulation. The zero-order hold circuit has a $(\sin x/x)$ response. The sampling images are the periodically repeated copies of the desired spectrum with frequency shifted by integer multiples of the sampling rate. The sampling images are considered to be interference terms and shall be eliminated by the reconstruction filters. In addition, these images

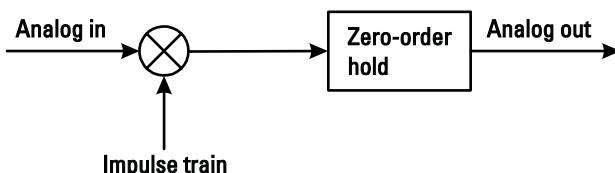


Figure 2.29 Effects of impulse train modulator and zero-order hold.

and interference terms are attenuated by the *sinc* response of the zero-order hold circuit in the D/A converter. Thus, the spectrum of the two component signals after passing the reconstruction filters can be described by [44]

$$S_{1,2}(f) = \sum_{k=-\infty}^{\infty} H_z(f) \cdot H_r(f) \cdot S_{1,2}(f + kf_s) \quad (2.75)$$

where $H_r(f)$ is the frequency response of the reconstruction filters, f_s is the sampling rate, and $H_z(f)$ is the “*sinc*” response of the zero-order hold circuit,

$$H_z(f) = \frac{1}{f_s} e^{j\pi f/f_s} \cdot \text{sinc}(\pi \frac{f}{f_s}) \quad (2.76)$$

Even though aliasing of the wideband quadrature signals could occur during the D/A conversion, the signal recombining still results in the complete cancellation of the wideband signals owing to the fact that they are ideally 180° out of phase. Now, the final output is

$$\begin{aligned} S_{\text{out}}(f) &= S_1(f) + S_2(f) \\ &= 2 \sum_{k=-\infty}^{\infty} H_z(f) \cdot H_r(f) \cdot s(f + kf_s) \end{aligned} \quad (2.77)$$

Note that the sampling images of the source signal are added in-phase, after being attenuated by the reconstruction filters and the “*sinc*” response. Figure 2.30 shows the spectrum for the upper amplifier branch signal $S_1(t)$. The reconstruction filter used is a Butterworth filter, and the sampling rate is $4\times$ oversampling. The cutoff frequency of the filter is fixed at 2.5 MHz. Figure 2.31 gives the output spectrum with three different filter orders for a CDMA IS-95 waveform.

In addition to the various filter types, the sampling images may be suppressed by increasing the sampling rate, increasing the filter order, or decreasing the filter cutoff frequency. However, note that the two component signals have a much wider bandwidth than the desired signal. As the filters cut more deeply into the signal at the higher frequencies, the envelope of the two component signals will start to ripple and introduce intermodulation through the saturated power amplifiers by the mechanisms we have already discussed. As usual, this intermodulation term may be mitigated with saturated- or switching-mode power amplifiers. Both the sampling images

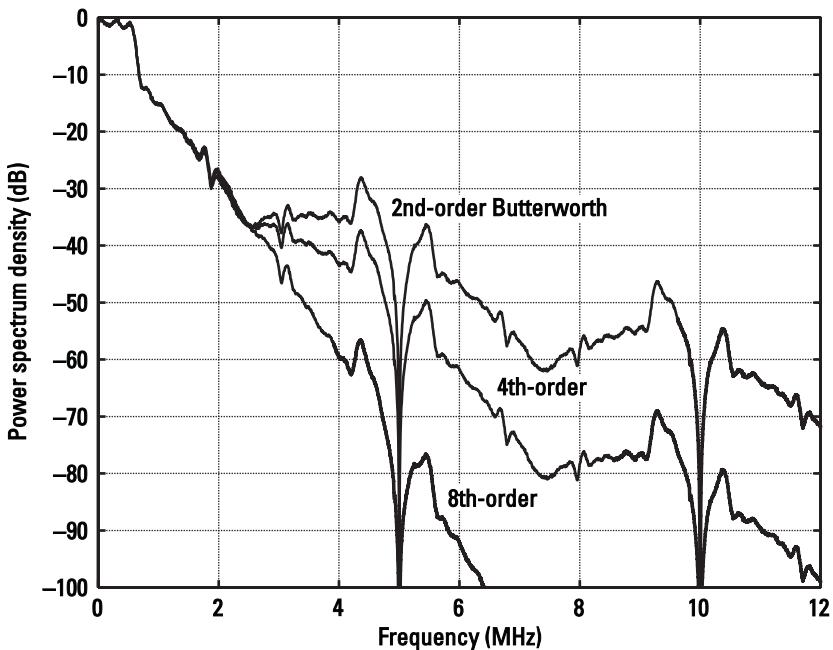


Figure 2.30 Simulated spectrum of the upper branch signal $s(t) + e(t)$ for CDMA IS-95 with differing Butterworth reconstruction filter order.

and intermodulation contribute to the out-of-band interference and should be kept below the specified ACI. It is generally a cumbersome task to find the optimum combination of sampling rate and filter cutoff frequency. A one-dimensional search method based on simulation was proposed in [44] and proved to be useful. This method assumes that an increase of the cutoff frequency decreases the intermodulation and that the residual spectrum near the sampling frequency is dominated by the sampling image. For a specified filter type and order, the optimal cutoff frequency was found by iterative calculation and simulation to suppress both the sampling image and intermodulation below the required ACI. Butterworth and Bessel filters were investigated; the former performs considerably better due to its sharper cutoff frequency.

2.9 Summary

The major factors contributing to the linearity degradation in outphasing power amplifier systems have been discussed. The path imbalance between

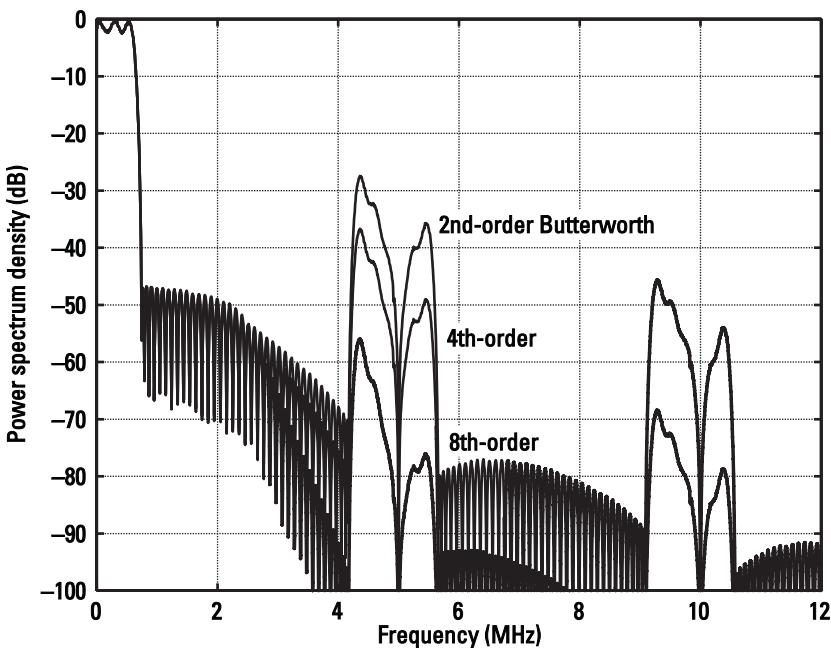


Figure 2.31 Simulated total output spectrum for CDMA IS-95 with differing Butterworth reconstruction filter orders.

the two power amplifier branches results in incomplete cancellation of the wideband quadrature signal and hence creates both out-of-band interference and inband distortion. The misalignment of the I/Q modulators includes the gain and phase error between the I and Q channels and the carrier leakage. These factors introduce amplitude variation to the two phase-modulated signals and create intermodulation through the AM-AM and AM-PM conversion of the nonlinear power amplifiers. The quadrature modulator errors may have more detrimental effects on the linearity performance of an outphasing system than on other linearization techniques. These error effects are more troublesome, since they cannot be compensated for as easily as the path imbalance effect. However, the quadrature modulator in an outphasing system is ideally producing a constant envelope output signal, and this fact may lead to a simpler and more effective approach to track and correct the quadrature modulator errors. The linearity performance is also degraded by the digital SCS due to quantization effects, sampling, and reconstruction filtering. The quantization noise of the source signal directly adds to the output as a constant background. The quantization of

the quadrature signal causes variation of the gain and phase characteristics of the nonlinear power amplifiers and hence creates interference terms. The effects of the reconstruction filters and sampling rate have also been addressed. The filter type, order, and cutoff frequency generally need to be traded off with the sampling rate to achieve the optimum performance. In the case of the envelope fluctuation induced by the quadrature modulator errors, reconstruction filtering, and DSP sampling, two limiters with sufficient bandwidth or two switching-mode power amplifiers may mitigate the linearity impairment created by the amplifier nonlinearities.

References

- [1] Proakis, J., *Digital Communications*, 2nd ed., New York: McGraw-Hill, 1989.
- [2] Feher, K., *Advanced Digital Communications: Systems and Signal Processing Techniques*, Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [3] Couch, L. W., *Digital and Analog Communication Systems*, 5th ed., Upper Saddle River, NJ: Prentice-Hall, 1997.
- [4] Feher, K., *Wireless Digital Communications: Modulations and Spread Spectrum Applications*, Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [5] Sevic, J. F., and M. B. Steer, "On the Significance of Envelope Peak-to-Average Ratio for Estimating the Spectral Regrowth of an RF/Microwave Power Amplifier," *IEEE Trans. MTT*, Vol. 48, No. 6, June 2000, pp. 1068–1071.
- [6] Shannon, C., "A Mathematical Theory of Communication," *Bell System Technical Journal*, Vol. 27, 1948, pp. 379–423.
- [7] Feynman, R., *Feynman's Lectures on Computation*, Reading, MA: Addison-Wesley, 1996.
- [8] Mehrotra, A., *GSM System Engineering*, Norwood, MA: Artech House, 1997.
- [9] EIA/TIA IS-95A, "Mobile Station—Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System," Telecommunications Industry Association, 1995.
- [10] Couch, L., and J. L. Walker, "A VHF LINC Amplifier," in *Proc. IEEE Southeastcon'82*, Destin, FL, April 1982, pp. 122–125.
- [11] Cox, D. C., "Linear Amplification with Nonlinear Components," *IEEE Trans. Commun.*, Vol. COM-22, December 1974, pp. 1942–1945.
- [12] Cox, D. C., and R. P. Leck, "Component Signal Separation and Recombination for Linear Amplification with Nonlinear Components," *IEEE Trans. Commun.*, Vol. COM-23, November 1975, pp. 1281–1287.

- [13] Cox, D. C., and R. P. Leck, "A VHF Implementation of a LINC Amplifier," *IEEE Trans. Commun.*, Vol. COM-24, September 1976, pp. 1018–1022.
- [14] Chan, K. Y., and A. Bateman, "Linear Modulators Based on RF Synthesis: Realization and Analysis," *IEEE Trans. Circ. Sys.*, Vol. 42, No. 6, 1995, pp. 321–333.
- [15] DaSilva, M. K., "Vector Locked Loop," U.S. Patent 5 105 168, April 14, 1992.
- [16] Okubo, N., et al., "Constant-Amplitude Wave Combination Type Amplifier," U.S. Patent 5 287 069, February 1994.
- [17] Hornak, T., and W. J. McFarland, "Power Amplifier Utilizing the Vector Addition of Two Constant Envelope Carriers," U.S. Patent 5 365 187, November 1994.
- [18] Rustako, A. J., and Y. S. Yeh, "A Wide-Band Phase-Feedback Inverse-Sine Phase Modulator with Application Toward a LINC Amplifier," *IEEE Trans. Commun.*, Vol. 24, October 1976, pp. 1139–1143.
- [19] Schemel, R. E., "Generating Arcsine(x) and Alternative Method for LINC," *Electron. Lett.*, Vol. 35, No. 10, May 1999, pp. 782–783.
- [20] Shi, B., and L. Sundstrom, "A 200-MHz IF BiCMOS Signal Component Separator for Linear LINC Transmitters," *IEEE JSSC*, Vol. 35, No. 7, July 2000, pp. 987–993.
- [21] Shi, B., and L. Sundstrom, "A Translinear-Based Chip for Linear LINC Transmitters," in *2000 Symposium on VLSI Circ. Dig. of Technical Papers*, Honolulu, HI, 2000, pp. 58–61.
- [22] Shi, B., and L. Sundstrom, "An IF CMOS Signal Component Separator Chip for LINC Transmitters," in *Proc. IEEE 2001 Custom Integrated Circuits Conference*, Piscataway, NJ, 2001, pp. 49–52.
- [23] Gilbert, B., "Translinear Circuits: A Proposed Classification," *Electron. Lett.*, Vol. 11, No. 1, 1975, pp. 14–16.
- [24] Seevinck, E., and R. J. Wiegerink, "Generalized Translinear Circuit Principle," *IEEE JSSC*, Vol. 26, No. 8, August 1991, pp. 1098–1102.
- [25] Bateman, A., R. J. Wilkinson, and J. D. Marvill, "The Application of Digital Signal Processing to Transmitter Linearisation," *IEEE 8th European Conf. on Electrotechnics*, 1988, pp. 64–67.
- [26] Hetzel, S. A., A. Bateman, and J. P. McGeehan, "LINC Transmitter," *Electron. Lett.*, Vol. 27, No. 10, May 1991, pp. 844–846.
- [27] Conradi, C. P., J. G. McRory, and R. H. Johnston, "Low-Memory Digital Signal Component Separator for LINC Transmitters," *Electron. Lett.*, Vol. 37, No. 7, March 2001, pp. 460–461.
- [28] Casadevall, F. J., and J. J. Olmos, "On the Behavior of the LINC Transmitter," in *Proc. 40th IEEE Veh. Technol. Conf.*, Orlando, FL, May 6–9, 1990, pp. 29–34.
- [29] Casadevall, F. J., and A. Valdovinos, "Performance Analysis of QAM Modulations Applied to the LINC Transmitter," *IEEE Trans. Veh. Technol.*, Vol. 42, No. 4, November 1993, pp. 399–406.

- [30] Sundstrom, L., "Automatic Adjustment of Gain and Phase Imbalances in LINC Transmitters," *Electr. Lett.*, Vol. 31, No. 3, 2 February 1995, pp. 155–156.
- [31] Roome, S. J., "Analysis of Quadrature Detectors Using Complex Envelope Notation," *Inst. Elec. Eng. Proc.*, Vol. 136, Pt. F, No. 2, April 1989, pp. 95–100.
- [32] Faulkner, M., and T. Mattsson, "Spectral Sensitivity of Power Amplifiers to Quadrature Modulator Misalignment," *IEEE Trans. Veh. Technol.*, Vol. 41, No. 4, November 1992, pp. 516–525.
- [33] Cavers, J. K., "The Effect of Quadrature Modulator and Demodulator Errors on Adaptive Digital Predistorters for Amplifier Linearization," *IEEE Trans. Veh. Technol.*, Vol. 46, No. 2, May 1997, pp. 456–466.
- [34] Gingell, M. J., "Single Sideband Modulation Using Sequence Asymmetric Polyphase Networks," *Electr. Comm.*, Vol. 48, 1973, pp. 21–25.
- [35] Koullias, I. A., et al., "A 900 MHz Transceiver Chip Set for Dual-Mode Cellular Radio Mobile Terminals," in *IEEE ISSCC Dig. of Tech. Papers*, February 1993, pp. 140–141.
- [36] Duncan, R., et al., "A 1-GHz Quadrature Sinusoidal Oscillator," *IEEE CICC Dig.*, 1995, pp. 91–94.
- [37] Faulkner, M., T. Mattsson, and W. Yates, "Automatic Adjustment of Quadrature Modulators," *Electron. Lett.*, Vol. 27, No. 3, January 1991, pp. 214–216.
- [38] Cavers, J. K., and M. W. Liao, "Adaptive Compensation for Imbalance and Offset Losses in Direct Conversion Transceivers," *IEEE Trans. Veh. Technol.*, Vol. 42, No. 4, November 1993, pp. 581–588.
- [39] Lohtia, A., P. Goud, and C. Englefield, "An Adaptive Digital Technique for Compensating for Analog Quadrature Modulator/Demodulator Impairments," in *Proc. IEEE Pac. Rim Conf.*, 1993, pp. 447–450.
- [40] Cavers, J. K., "New Methods for Adaption of Quadrature Modulators and Demodulators in Amplifier Linearization Circuits," *IEEE Trans. Veh. Technol.*, Vol. 46, No. 3, August 1997, pp. 707–716.
- [41] Sundstrom, L., "Spectral Sensitivity of LINC Transmitters to Quadrature Modulator Misalignments," *IEEE Trans. Veh. Technol.*, Vol. 49, No. 4, July 2000, pp. 1474–1487.
- [42] Oppenheim, A. V., and R. W. Schafer, *Discrete-Time Signal Processing*, Englewood Cliffs, NJ: Prentice Hall, 1989.
- [43] Sundstrom, L., "The Effect of Quantization in a Digital Signal Component Separator for LINC Transmitters," *IEEE Trans. Veh. Technol.*, Vol. 45, No. 2, May 1996, pp. 346–352.
- [44] Sundstrom, L., "Effects of Reconstruction Filters and Sampling Rate for a Digital Signal Component Separator on LINC Transmitter Performance," *Electron. Lett.*, Vol. 31, No. 14, July 1995, pp. 1124–1125.

3

Path Mismatch Reduction Techniques for Outphasing Amplifiers

3.1 Introduction

Linear modulations such as M-PSK and M-QAM are desired for modern digital wireless communications because of their superior spectral efficiency, but they require linear power amplification at transmission. On the other hand, limited battery capacity imposes primary restrictions on the power consumption of mobile handsets, and even base station power amplifiers have limited dc power availability. It is typically the final power amplifier stage in the transmitter that constitutes this real challenge for the designer—the power efficiency and linearity need to be carefully traded off since they generally require different tuning and cannot be maximized simultaneously. Conventional linear amplifier design techniques, like load pull, entail a great deal of effort to optimize the amplifier performance due to the high cost of power transistors. The outphasing power amplifier takes advantage of two nonlinear amplifiers to achieve linear amplification and hence eliminates this trade-off and greatly simplifies design procedures in this respect. Moreover, the inherent characteristic of this architecture allows power amplifiers to continuously operate at their peak power efficiency and potentially improves the overall efficiency of the system.

The outphasing amplifier, differing from other linearization techniques, achieves linear amplification with a pair of “matched” nonlinear power amplifiers. Instead of the individual amplifier linearity, the matching condition turns out to be the major concern, provided that the amplifier has sufficient bandwidth to accommodate the phase-modulated component

signals. Among the factors that contribute to the system linearity degradation in an outphasing amplifier, the RF path mismatch between two amplifier paths is the major mismatch source. Unfortunately, this architecture demands an extremely tight tolerance on the matching condition for the acceptably small out-of-band rejection. This problem has been analyzed [1–3], and typical requirements for most practical applications are approximately 0.1 to 0.5 dB in gain matching and 0.4 to 2° in phase matching. Furthermore, the gain and phase characteristics of the power amplifiers are highly variable due to thermal drift, component aging, and channel transition. Therefore, the matching condition must be maintained free of environmental variations. This is nearly impossible to achieve in most practical situations and has been the primary reason that prevents outphasing approach from the practical applications. Attempts have been made to correct for the path imbalance through some kind of feedback approaches.

This chapter first introduces the path mismatch correction schemes based on training vectors. With these methods, the gain and phase imbalance between two amplifier branches are characterized by a set of tests, and then compensated by introducing a predistorted term. These schemes are simple and effective, but interrupt regular data transmission. Correction techniques that can operate in the *background* are then described. These schemes allow simultaneous data transmission and calibration and therefore are applicable to any communication standards. For broadband and multicarrier applications such as in a base station, a high degree of similarity between two amplifier channels is demanded throughout the entire bandwidth—the simple gain and phase calibration and compensation technique at one frequency is not adequate, and channel equalization must be employed to balance the responses of two branches for the whole channel. This generalized calibration scheme is also discussed. Finally, VCO-derived syntheses techniques are reviewed, including the CALLUM and the VLL. These two methods were proposed as alternative approaches to realize signal component separation and are able to automatically correct the path imbalance. The latter characteristic represents a unique advantage over other correction techniques, which makes them particularly suitable for low-power and low-cost integrated circuit implementation.

3.2 Correction Schemes Based on Training Vectors

These correction techniques employ efficient training vectors that are not consistent with the transmitted data sequences and therefore require

a dedicated time period during regular data transmission for the calibration and continuous adjustment of the imbalance information. However, the rate of update on the imbalance can be much less frequent than the data rate, since the environmental fluctuations that cause it are generally much slower processes. These schemes could be directly implemented in wireless systems that employ time-division duplex (TDD) [e.g., digital European cordless telephone (DECT) and PHSs, in which calibration could be processed when the transmitter is “off”]. This is not true for continuous transmission systems utilizing FDMA or CDMA. For TDMA systems, a certain short time slot might be reserved for all users for calibration purpose only; utilizing other users’ time slots for calibration at a specific frequency outside of the band is an alternative way. Another possible approach for TDMA or CDMA systems might be to embed the calibration code at the beginning of each individual data packet, where the data packet is much longer than the calibration code.

3.2.1 Baseband Preconditioning of Path Mismatch Errors

A baseband preconditioning technique was proposed in [4, 5]. This method is based on the RF power measurements at the power summing ports. The quadrature modulator errors are corrected in advance of the path imbalance, if necessary. The measurement and compensation of the quadrature modulator errors are similar to the methods described in Chapter 2, by use of two power detectors for each modulator as shown in Figure 3.1. The LO leakage is first corrected by zeroing the baseband I/Q inputs and adjusting the dc offsets such that the detected RF output is minimized. The correction is made iteratively, and the power detector does not need to be linear or calibrated [6]. The gain error is then abstracted by applying the I/Q vector (1,0) and (0,1) successively and comparing the detected power. Finally, applying test vector (1,1) and (1,−1) determines the phase error δ

$$\delta \approx \frac{1}{2} \left[\frac{P_{(1,1)}}{P_{(1,-1)}} - 1 \right] \quad (3.1)$$

The compensation of the quadrature modulator error can be easily achieved by adjusting the dc offset, amplitude, and phase of the baseband I/Q signal inside the DSP. A similar strategy can be applied to the measurement of path imbalance. Specifically, the gain imbalance is determined by separately applying the I/Q vector (1,0) to one of the two branches while disabling the other, and vice versa. Then the phase imbalance is determined

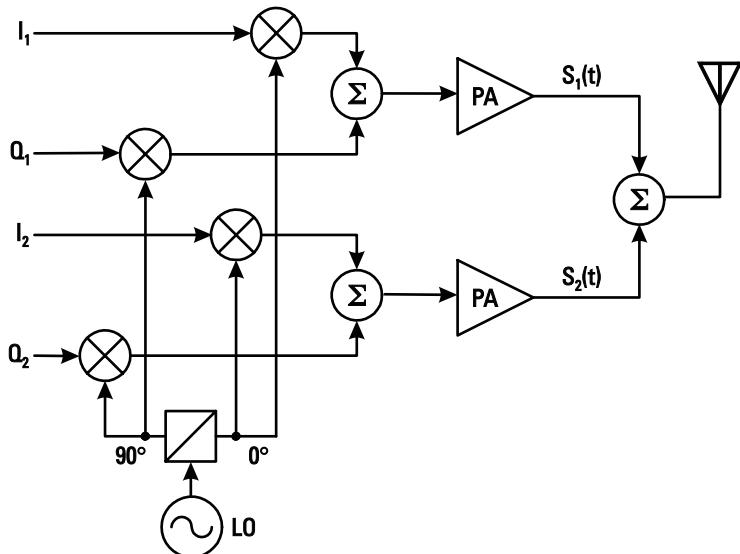


Figure 3.1 Block diagram for baseband preconditioning [4].

by taking the power measurement of test vector $(1,0)$ for the upper branch while disabling the lower one for P_{ref} , and $(-1,0)$ for the lower one for P_0 .

$$\cos \Delta\phi = 1 - \frac{1}{2} \frac{P_0}{P_{\text{ref}}} \quad (3.2)$$

The compensation of the phase imbalance is realized by simply phase-shifting one of the two component signals. This can be accomplished in the RF/LO with a phase shifter, or inside the baseband DSP circuit. The compensation of the gain imbalance cannot be done in the baseband, since highly nonlinear power amplifiers are preferred, and it is difficult to adjust the gain of such a nonlinear power amplifier while maintaining its high power efficiency. For the saturated amplifiers, the optimum approach might be to “back-off” one power amplifier with the highest output power until the two amplifiers have equal power. This would allow the power amplifier with the smaller output power to operate at peak efficiency and direct correction in baseband. However, this approach is not applicable to switching-mode power amplifiers. A more general solution applicable to both cases is to use the dc–dc converter as the power supply of the nonlinear amplifiers. The gain imbalance is compensated for by adjusting the supply voltage of the power amplifier, and the amplifier output power varies approximately proportional to the square of supply voltage.

3.2.2 Foreground Calibration Algorithm of Path Mismatch Errors

The foreground calibration algorithm makes use of the standard of amplitude and phase produced by the baseband DSP to calibrate the amplifiers [7]. As illustrated in Figure 3.2, a feedback loop is added to the standard outphasing amplifier system to characterize the gain and phase imbalance. A directional coupler withdraws a small portion of the system output and applies it to a downconversion mixer, driven by the same LO as the quadrature modulators. After lowpass filtering, the signal is then analog-to-digital (A/D) converted and sent to the baseband DSP. The DSP determines the gain and phase imbalance and eliminates the error effects by introducing a predistortion term inside the DSP. The following analysis illustrates the derivation of this algorithm.

3.2.2.1 Foreground Calibration Algorithm

Suppose that G_0 is the amplifier gain, ϕ_0 is the path delay, and the gain and phase imbalance of the lower amplifier with respect to the upper one are ΔG and $\Delta\phi$, respectively. Then, the combined amplifier output will be

$$\begin{aligned} S(t) &= S_1(t) + S_2(t) \\ &= G_0 r_{\max} \cos [\omega_c t + \theta(t) - \psi(t) + \phi_0] \\ &\quad + (G_0 + \Delta G) r_{\max} \cos [\omega_c t + \theta(t) + \psi(t) + \phi_0 + \Delta\phi] \end{aligned}$$

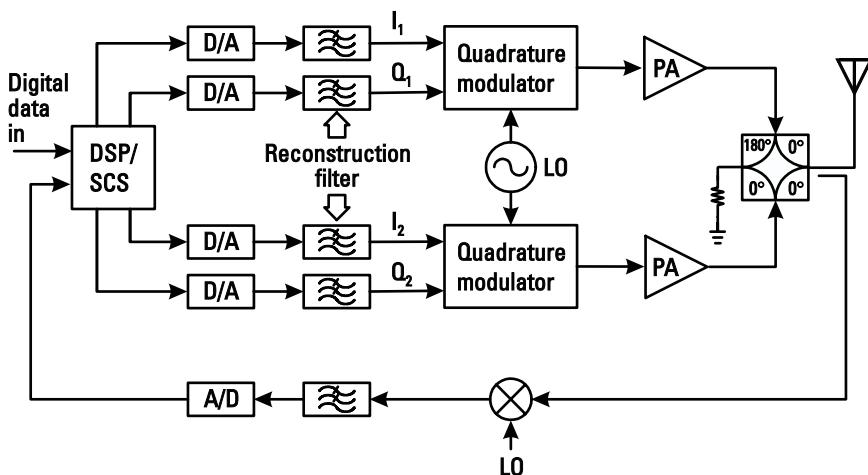


Figure 3.2 Outphasing amplifier diagram with foreground calibration loop [7].

$$\begin{aligned}
&= 2G_0 \cos\left(\frac{1}{2}\Delta\phi\right) r(t) \cos[\omega_c t + \theta(t) + \phi_0 + \frac{1}{2}\Delta\phi] \\
&\quad + \Delta G r(t) \cos[\omega_c t + \theta(t) + \phi_0 + \Delta\phi] \\
&\quad - 2G_0 \sqrt{r_{\max}^2 - r^2(t)} \sin\left(\frac{1}{2}\Delta\phi\right) \cos\left[\omega_c t + \theta(t) + \phi_0 + \frac{1}{2}\Delta\phi\right] \\
&\quad - \Delta G \sqrt{r_{\max}^2 - r^2(t)} \sin[\omega_c t + \theta(t) + \phi_0 + \Delta\phi]
\end{aligned} \tag{3.3}$$

where ω_c is the carrier frequency. The last expression is obtained by first combining the phase distortion terms and then abstracting $\psi(t)$ out of the phase angle. The spectrum regrowth stems from the amplitude and phase distortion of the summed signal. The first two terms in (3.3) contribute to the desired signal, and the last two terms add to the out-of-band spectrum, as a consequence of incomplete cancellation of the wideband quadrature signal $e(t)$. The ACI is related to the ratio of the last two terms to the first two terms, which is proportional to the gain and phase imbalance, as shown in Chapter 2.

The variables $r(t)$ and $\theta(t)$ —the amplitude and phase of the baseband digital input modulated by the SCS—determine the relation between the four terms in (3.3). The correct choice of $r(t)$ and $\theta(t)$ provides a mechanism to abstract the gain and phase error information from the distorted output signal. Specifically, four different combinations of $r(t)$ and $\theta(t)$ are sufficient to obtain the gain and phase imbalance from the combined output. The foreground algorithm takes the approach of setting these two variables to extreme cases— $r(t) = r_{\max}$ or 0, and $\theta(t) = 0$ or $\pi/2$. Equivalently, we can set the baseband I/Q components of two signal vectors to be 0 or $\pm r_{\max}$, which is straightforward to realize with the DSP. The correction algorithm consists of the successive generation and measurement of four calibration signals. First, we set the amplitude of the baseband input signal to the maximum allowable level of the DSP/SCS [i.e., $r(t) = r_{\max}$]. Then setting $\theta(t) = 0$ to get S_0 , and $\theta(t) = \pi/2$ to get S_p . The results from (3.3) after downconversion are

$$\begin{aligned}
S_0 &= G'_L r_{\max} \cos\left(\frac{1}{2}\Delta\phi\right) \cos\left(\phi_L + \frac{1}{2}\Delta\phi\right) \\
&\quad - \frac{1}{2} \frac{\Delta G}{G_0} G_L r_{\max} \sin\left(\frac{1}{2}\Delta\phi\right) \sin\left(\phi_L + \frac{1}{2}\Delta\phi\right) \\
&\approx G'_L r_{\max} \cos\left(\phi_L + \frac{1}{2}\Delta\phi\right)
\end{aligned} \tag{3.4}$$

$$\begin{aligned}
S_p &= -G'_L r_{\max} \cos\left(\frac{1}{2} \Delta\phi\right) \sin\left(\phi_L + \frac{1}{2} \Delta\phi\right) \\
&\quad - \frac{1}{2} \frac{\Delta G}{G_0} G_L r_{\max} \sin\left(\frac{1}{2} \Delta\phi\right) \cos\left(\phi_L + \frac{1}{2} \Delta\phi\right) \\
&\approx -G'_L r_{\max} \sin\left(\phi_L + \frac{1}{2} \Delta\phi\right)
\end{aligned} \tag{3.5}$$

where

$$G'_L = \left(1 + \frac{\Delta G}{2G_0}\right) G_L \tag{3.6}$$

and G_L is the effective loop gain; ϕ_L consists of the loop delay and phase shift introduced by the mixer. We then set the amplitude of the input baseband signal to zero [*i.e.*, $r(t) = 0$]. Now, the first two terms in (3.3) are removed. As before, we set $\theta(t) = 0$ to obtain S_a

$$\begin{aligned}
S_a &= -G_L r_{\max} \sin\left(\frac{1}{2} \Delta\phi\right) \cos\left(\phi_L + \frac{1}{2} \Delta\phi\right) - \frac{1}{2} \frac{\Delta G}{G_0} G_L r_{\max} \sin(\phi_L + \Delta\phi) \\
&\approx -\frac{1}{2} \Delta\phi G_L r_{\max} \cos\left(\phi_L + \frac{1}{2} \Delta\phi\right) - \frac{1}{2} \frac{\Delta G}{G_0} G_L r_{\max} \sin\left(\phi_L + \frac{1}{2} \Delta\phi\right)
\end{aligned} \tag{3.7}$$

and set $\theta(t) = \pi/2$, to obtain S_b

$$\begin{aligned}
S_b &= G_L r_{\max} \sin\left(\frac{1}{2} \Delta\phi\right) \sin\left(\phi_L + \frac{1}{2} \Delta\phi\right) - \frac{1}{2} \frac{\Delta G}{G_0} G_L r_{\max} \cos(\phi_L + \Delta\phi) \\
&\approx \frac{1}{2} \Delta\phi G_L r_{\max} \sin\left(\phi_L + \frac{1}{2} \Delta\phi\right) - \frac{1}{2} \frac{\Delta G}{G_0} G_L r_{\max} \cos\left(\phi_L + \frac{1}{2} \Delta\phi\right)
\end{aligned} \tag{3.8}$$

Making use of (3.4), (3.5), (3.7), and (3.8), S_a and S_b can be written in matrix form as

$$\begin{pmatrix} S_a \\ S_b \end{pmatrix} \approx -\frac{1}{2} \begin{pmatrix} -S_p & S_0 \\ S_0 & S_p \end{pmatrix} \begin{pmatrix} \Delta G/G_0 \\ \Delta\phi \end{pmatrix} \tag{3.9}$$

Solving (3.9) for $\Delta G/G_0$ and $\Delta\phi$ yields

$$\frac{\Delta G}{G_0} \approx \frac{1}{P_L} (S_p S_a - S_0 S_b) \tag{3.10}$$

$$\Delta\phi \approx -\frac{1}{P_L} (S_0 S_a + S_p S_b) \quad (3.11)$$

where P_L is defined by

$$P_L = \frac{1}{2} G_L^2 r_{\max}^2 \quad (3.12)$$

Note that if two amplifier branches are perfectly matched, $G_L r_{\max}$ actually corresponds to the maximum signal amplitude detected by the DSP/SCS at the end of correction loop; hence P_L might be referred to as the “average” power level normalized to a 1Ω characteristic impedance during calibration. The quantity P_L can be estimated by

$$P_L \approx \frac{1}{2} (S_0^2 + S_p^2) \quad (3.13)$$

Equation (3.9) indicates that the gain and phase imbalance are solely determined by the resulting output of four calibration signals S_0 , S_p , S_a , and S_b . The relationship between them, as well as between P_L and ϕ_L , are best illustrated by Figure 3.3 in the case of small gain and phase imbalance, where

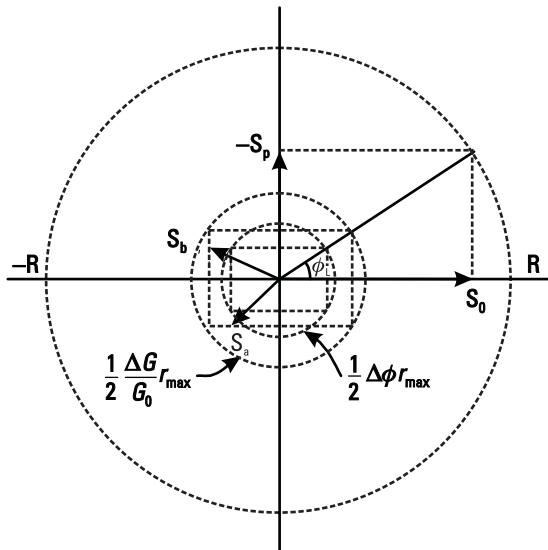


Figure 3.3 Relationship between path imbalance and calibration signals ($R = G_L r_{\max}$).

S_a and S_b are the linear combinations of S_0 and S_p scaled by the gain and phase imbalance. Observed from the antenna, S_a and S_b are near zero-power-level signals, and S_0 and S_p are full-power-level signals. If two amplifier branches are perfectly matched, S_a and S_b would be zero.

The approximations in (3.4), (3.5), and (3.9) give rise to a certain amount of estimation error for the measurement of the gain and phase imbalance. Note that as long as the gain and phase error are close to zero, there will be negligible estimation error in (3.10). Therefore the estimate and compensation of gain and phase imbalances could be iterative; with several iterations, the gain and phase imbalances are able to converge to an arbitrarily low level. Also note that the measured gain and phase imbalance are proportional to $1/P_L$, and that suggests that a precise determination of P_L is not essential and that we may need to estimate this parameter only once at the beginning of the calibration. To see how the path imbalance affects the algorithm convergence, let's acquire the accurate expressions of S_0 , S_p , S_a , and S_b from (3.3), and substitute them and the definition of P_L into (3.10). After applying a trigonometric expansion, it is not difficult to obtain the following expression for the gain imbalance

$$\left(\frac{\Delta G}{G_0}\right)_{\text{est}} = \left(1 + \frac{\Delta G}{2G_0}\right)\left(\frac{\Delta G}{G_0}\right)_{\text{act}} \quad (3.14)$$

where “est” and “act” indicate estimate and actual, respectively. Note that no approximation has been made to derive the above expression. A similar result can be obtained for the phase imbalance

$$\begin{aligned} (\Delta\phi)_{\text{est}} &= \left(1 + \frac{\Delta G}{G_0}\right) \sin(\Delta\phi)_{\text{act}} \\ &\approx \left(1 + \frac{\Delta G}{G_0}\right)(\Delta\phi)_{\text{act}} \end{aligned} \quad (3.15)$$

The only approximation here is the Taylor expansion of the sine function, which is accurate up to the second order of the phase imbalance and is negligible compared to the first-order estimation error introduced by the gain imbalance. Equations (3.14) and (3.15) conclude that after each iteration and compensation, the actual gain imbalance decreases by a factor of $\Delta G/2G_0$, while the actual phase imbalance decreases by a factor of $\Delta G/G_0$. Note that the iteration convergence is independent of the phase imbalance. The actual gain phase imbalance at the end of the n th iteration and compensation are given by

$$\left. \frac{\Delta G}{G_0} \right|_{\max, n} \approx 2 \left(\frac{\Delta G}{2G_0} \right)^{(2^n)} \quad (3.16)$$

$$\left. \Delta \phi \right|_{\max, n} \approx 2\Delta\phi \left(\frac{\Delta G}{2G_0} \right)^{(2^n-1)} \quad (3.17)$$

In the ideal case, the gain and phase imbalance would be no more than $\pm 2\%$ and $\pm 2.3^\circ$ after the first compensation, and less than 0.1% and 0.3° after the second iteration, assuming that the initial gain and phase imbalance are 20% and 11.5° , respectively. In other words, at most, three iterations are needed for the calibration to converge.

3.2.2.2 Practical Considerations

Perfectly balanced quadrature modulators are assumed in this architecture, and that leads to an important concern. The quadrature errors create a residue in the adjacent channels and this effect has been analyzed in Section 2.6.2. Now the outphasing amplifier output in the presence of quadrature errors—the amplitude error $g_{1,2}$ and phase error $\delta_{1,2}$ of each I/Q modulator—is described by,

$$\begin{aligned} S(t) &= S_1(t) + S_2(t) \\ &\approx (3.3) \\ &- \frac{1}{2} G_0(g_1 + g_2)r(t) \sin[\theta(t)] \sin(\omega_c t + \phi_0) \\ &- \frac{1}{2} G_0(\delta_1 + \delta_2)r(t) \sin[\theta(t)] \cos(\omega_c t + \phi_0) \\ &+ \frac{1}{2} G_0(g_1 - g_2)\sqrt{r_{\max}^2 - a^2(t)} \cos[\theta(t)] \sin(\omega_c t + \phi_0) \\ &+ \frac{1}{2} G_0(\delta_1 - \delta_2)\sqrt{r_{\max}^2 - a^2(t)} \cos[\theta(t)] \cos(\omega_c t + \phi_0) \end{aligned} \quad (3.18)$$

The $r(t)$ terms are narrowband and introduce inband distortion, while the $\sqrt{r_{\max}^2 - a^2(t)}$ terms are wideband and create out-of-band spectrum. The out-of-band spectrum can only be suppressed by matching the two I/Q modulators, not by adjusting the gain and phase delay of the amplifier branches.

The presence of the quadrature errors degrades the measurement accuracy of the gain and phase imbalance. Following procedures similar to those used to derive the signal after the lowpass filter (LPF), it is not difficult

to obtain the following expression for estimated gain and phase imbalance by using (3.10) and (3.18), and the definitions of S_0 , S_p , S_a , and S_b —that is,

$$\left(\frac{\Delta G}{G_0}\right)_{\text{est}} = \left(\frac{\Delta G}{G_0}\right)_{\text{act}} - (g_1 - g_2) \sin^2 \phi_L - (\delta_1 - \delta_2) \sin \phi_L \cos \phi_L \quad (3.19)$$

$$(\Delta\phi)_{\text{est}} = (\Delta\phi)_{\text{act}} - (\delta_1 - \delta_2) \cos^2 \phi_L - (g_1 - g_2) \sin \phi_L \cos \phi_L \quad (3.20)$$

The measured gain and phase imbalance are constantly offset by a certain small amount, depending on the quadrature errors and the phase delay of the correction loop. As a result, the gain and phase errors will finally reach values of

$$\left.\frac{\Delta G}{G_0}\right|_{n \rightarrow \infty} = (g_1 - g_2) \sin^2 \phi_L + (\delta_1 - \delta_2) \sin \phi_L \cos \phi_L \quad (3.21)$$

$$\left.\Delta\phi\right|_{n \rightarrow \infty} = (\delta_1 - \delta_2) \cos^2 \phi_L + (g_1 - g_2) \sin \phi_L \cos \phi_L \quad (3.22)$$

Practically, many imperfection factors contribute to the quadrature errors (e.g., LO leakage and temperature variation, not only the mismatching of amplitude and phase), and these two values are statistically averaged effects and highly variable with environment. After a few iterations, the measured imbalance “appears” close to zero, while the actual imbalance stays around the same order as the quadrature errors and varies slightly after each iteration. In other words, the quadrature errors set an upper limit for the overall performance of the outphasing amplifier. Fortunately, highly accurate quadrature modulators are now routinely available for upconversion and downconversion applications, with gain and phase accuracy well within the requirements of this proposed algorithm.

In practice, the dc offset of the mixer, LPF, and A/D converter add to the measured calibration signals as a constant background, and that degrades the measurement accuracy and algorithm convergence. To eliminate this effect, we may add an extra calibration signal S_c to determine the dc offset. S_c may be chosen to be a signal that is 180° out-of-phase with any of the four calibration signals [e.g., we choose S_c to be 180° out-of-phase with S_a , by setting $r(t) = 0$ and $\theta(t) = \pi$, i.e. $I_1 = I_2 = 0$, $Q_1 = r_{\max}$, and $Q_2 = -r_{\max}$]. The dc offset is easily determined by

$$V_{\text{offset}} = \frac{1}{2} (S_a + S_c) \quad (3.23)$$

Alternatively, the dc offset may be directly measured by disabling the transmitted signal.

In a typical mobile station, a duplexer is employed to isolate the transmitter from the receiver in order to prevent interference between the transmission band and receiving band. It is important to minimize the out-of-band spectral power during calibration since the duplexer cannot prevent the interference to adjacent channels. From (3.3), it is clear that during the calibration when amplitude $r(t)$ is fixed and $\theta(t)$ is swept, the amplifier output is a phase modulated signal. The transmission bandwidth— B_T —of the PM signals is given by Carson's rule [8],

$$\begin{aligned} B_T &= 2(\beta + 1)B \\ &= 5.1B \end{aligned} \quad (3.24)$$

where β is the phase modulating index, in this case $\pi/2$, and B is the bandwidth of the modulating signal. As we fix $\theta(t)$ and vary the amplitude $r(t)$, the output is a combination of four double-sideband suppressed carrier (DSB-SC) signals. For this amplitude modulation, the output power spectrum is the linear translation of the modulating signal. Therefore, the bandwidth of the amplifier output can be well controlled during correction.

The quantization error of the A/D converter may limit the performance of the correction algorithm. By definition, the dynamic range is the ratio of the maximum detectable signal level to the minimum detectable level. For a $(B + 1)$ -bit A/D converter, we have

$$\Delta = \frac{X_m}{2^B} \quad (3.25)$$

where Δ and X_m are the step size and the full-scale level of the A/D converter, respectively. The maximum detectable signal level is $X_m = G_L r_{\max}$. Since the quantization error is not more than $\Delta/2$, the worst case estimate of $\Delta\phi$ from (3.10) may be written as

$$\begin{aligned} (\Delta\phi)_{\text{est}} &= -\frac{2}{(G_L r_{\max})^2} \left[\left(S_0 + \frac{\Delta}{2} \right) \left(S_a + \frac{\Delta}{2} \right) + \left(S_p + \frac{\Delta}{2} \right) \left(S_b + \frac{\Delta}{2} \right) \right] \\ &\approx (\Delta\phi)_{\text{act}} + \frac{\sqrt{2}}{2^B} \sin\left(\phi + \frac{1}{2}\Delta\phi - \frac{\pi}{4}\right) \end{aligned} \quad (3.26)$$

The estimation error $|\Delta\phi|_{\text{err}}$ of the phase imbalance is bounded by

$$|\Delta\phi|_{\text{err}} \leq \frac{\sqrt{2}}{2^B} \quad (3.27)$$

A similar result applies to the estimation error $|\Delta G/G_0|_{\text{err}}$ of the gain imbalance from (3.10),

$$\left| \frac{\Delta G}{G_0} \right|_{\text{err}} \leq \frac{\sqrt{2}}{2^B} \quad (3.28)$$

Then the required word length of the A/D converter is calculated to be

$$B + 1 = 1.5 - \frac{\log_{10} \delta}{\log_{10} 2} \quad (3.29)$$

with δ being the smaller value between the allowable gain and phase error. As an example, correction of I/Q modulator quadrature errors of 3.5% (0.3 dB) in amplitude and 2° in phase would require a $B + 1 = 6.3$, thus 7-bit, A/D converter. Considering the refresh rate of the calibration iteration, which might be tens of kilohertz, this is easily achievable with modern A/D converter technology.

Figure 3.4(a) displays an example of the calibration signal waveforms I1 and Q1. Waveforms I2 and Q2 are not shown here; they are time-shifted versions of Q1 and I1. The dashed lines indicate the moment when the calibration signal samples are taken. In between the dashed lines are the transitions that must be carefully designed to minimize out-of-band radiation. No out-of-band spurs above -55 dBc were observed in our experiment during calibration. The LPF transient effect is not crucial in this example, with transient time 2 to 3 orders less than the iteration period. This fact is demonstrated by Figure 3.4(b), in which the traces are smooth and no ripples are observed.

Note that an extra calibration signal S_C is added to determine the dc offset of the LPF and analog-to-digital converter (ADC). The calibration was accomplished within two or three iterations. Each iteration took 0.12 ms and can be further reduced if necessary. Figure 3.4(b) compares traces before and after calibration. For a perfectly matched outphasing system, S_a , S_b , and S_c should remain zero—the constant dc offset in this case, which is consistent with the experiment. Since CDMA IS-95 mobile terminals transmit signals in bursts, the calibration could be taken very shortly before real data transmission, with two zero-RF-power-level signals (S_a and S_b) followed by two full-RF-power-level signals (S_0 and S_p) at the beginning of every few bursts.

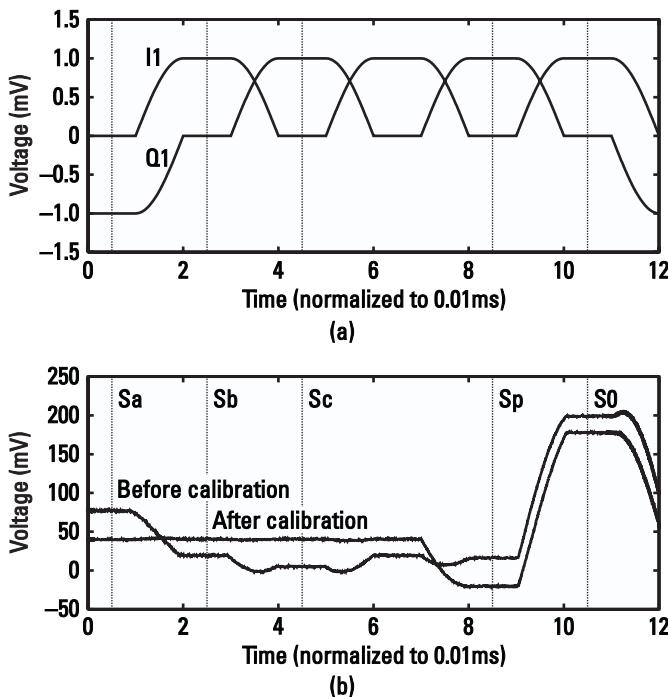


Figure 3.4 (a) Calibration signal waveforms I1/Q1, and (b) lowpass filtered calibration signals before and after correction.

Figure 3.5 displays the output power spectrum of the amplifier system. The upper amplifier operates at gain compressed by 2 dB, as a fairly nonlinear amplifier. To adjust the power of the bottom branch to compensate the gain imbalance, the bottom amplifier operates 3 dB backed off from the 1-dB compression point in the linear operation region. The maximum capable output power of this system is 31 dBm. Without correction (upper curve), the ACI is around -27 dB, while with correction (lower curve) the out-of-band spectrum is suppressed efficiently below -38 dBc. The measured gain and phase imbalance prior to calibration are 0.5 dB and 12° , respectively. The relatively large phase imbalance comes from the fact that two power amplifiers are driven at different power levels, with one saturated and the other linearly operated. The suppression of the secondary spectral emission is less successful, due to the relatively large cutoff frequency of the reconstruction filters and their mismatches. Considering that the modulators used have a nominal 2° phase error and 0.3-dB gain error, the -38 dBc ACI is an outstanding result.

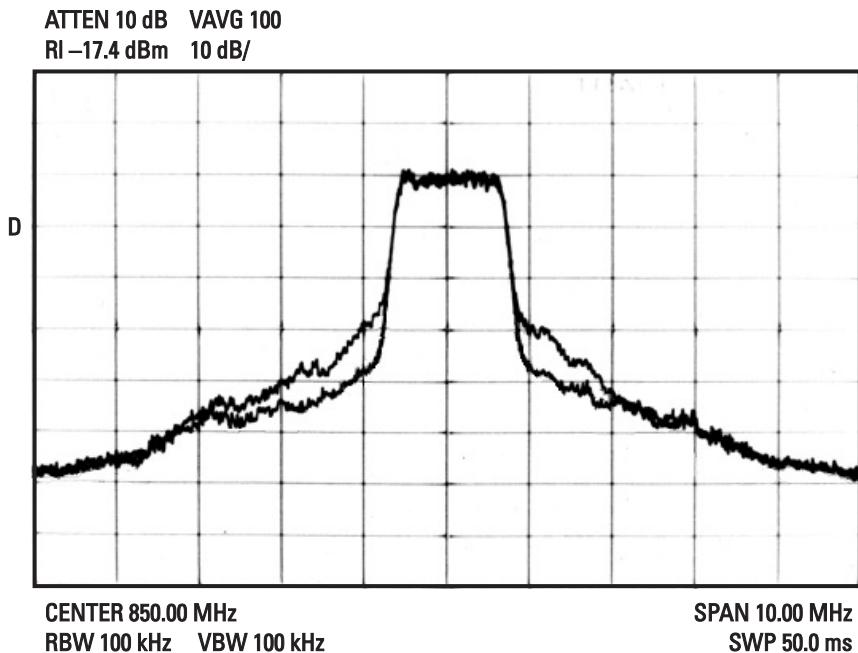


Figure 3.5 Measured output spectra for CDMA IS-95 with and without calibration using foreground correction scheme [9].

3.3 Path Mismatch Error Correction Schemes Transparent to Data Transmission

3.3.1 Phase-Only Correction Approach

A “phase-only” correction was proposed in [3]. In this method, the phase imbalance is detected by multiplying two power amplifier outputs, as illustrated in Figure 3.6. The multiplier output is lowpass filtered, differentiated, and fed to a comparator. Since a 90° hybrid is used as a power combiner, an extra 90° phase shift is added between the two component signals. In the presence of the phase imbalance δ between the two amplifier outputs, the differentiated output can be described by

$$S_{\text{LPF}}(t) = \cos[2\psi(t) + \delta] \cdot \psi'(t) \quad (3.30)$$

where

$$\cos \psi(t) = \frac{r(t)}{r_{\max}} \quad (3.31)$$

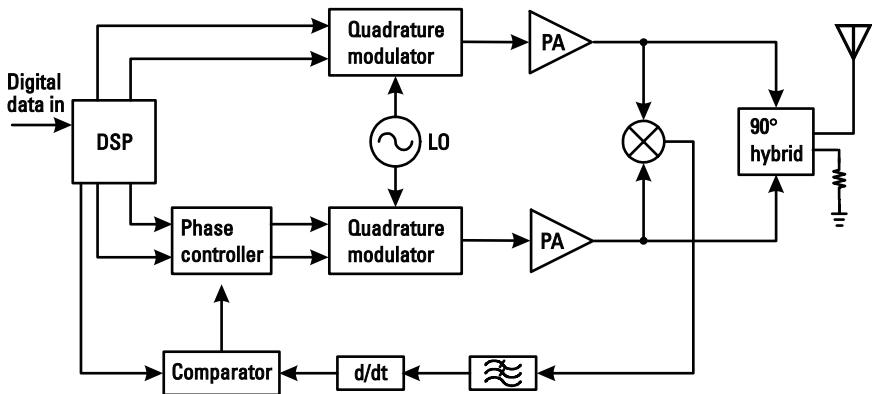


Figure 3.6 Outphasing amplifier with phase imbalance correction [3].

The DSP circuit generates a trigger pulse when two component signals are 180° out of phase, corresponding to $\psi(t) = 45^\circ$. The trigger time is then compared to the zero-crossing time of the detected signal, and the difference is used to guide a one-dimensional search for the phase imbalance. One branch signal phase is controlled by adding or subtracting a certain phase increment, and the phase imbalance is compensated. A variation of this approach is to use a 180° hybrid without the differentiator, although the differentiation operation helps to eliminate the effect of filter dc offset. Note that the zero-crossing time of the detected signal is used to guide the compensation of the phase imbalance, this method is free from the change of the amplifier output power. However, since the phase imbalance is determined by multiplying the two amplifiers' outputs, any imbalance after the power amplifiers—like the hybrid—is ignored. In addition, careful design is required to prevent the additional phase imbalance introduced by the measurement circuit.

3.3.2 Simplex Search Algorithm Correction

A simplex search algorithm was proposed in [10] to correct for both gain and phase imbalance. Figure 3.7 shows the schematic diagram of one implementation of this approach. The combined output from two power amplifiers is downconverted and fed to a bandpass filter (BPF) to recover the out-of-band emission that carries the gain and phase imbalance information. The magnitude of the filtered signal is integrated over a finite number of symbols. The integration guides a search algorithm to find the optimal gain and phase correction by minimizing the out-of-band power. As discussed in

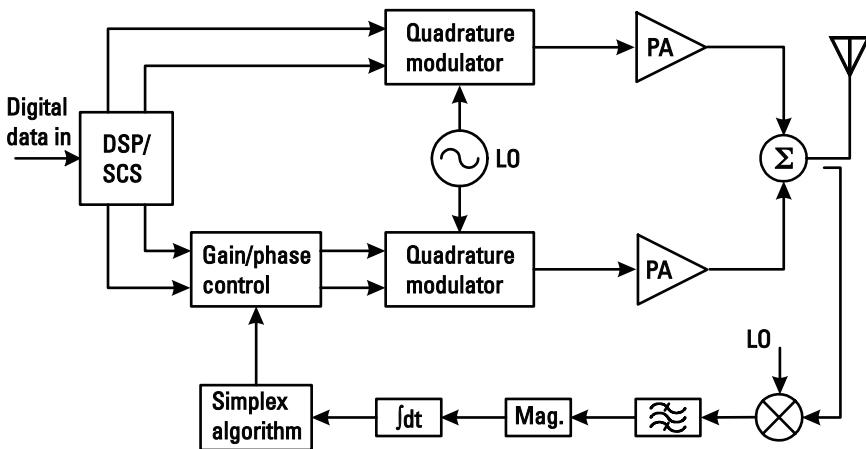


Figure 3.7 Block diagram for outphasing amplifier with simplex search algorithm [10].

Chapter 2, the out-of-band power of the combined output is proportional to the path imbalance

$$P_{\text{out-of-band}} \propto \left(\frac{\Delta G}{G_0} \right)^2 + \Delta\phi^2 \quad (3.32)$$

A straightforward optimization to find the gain and phase imbalance is through the *alternating variable* method. In this method, each variable is iteratively changed to reduce the objective function—out-of-band power in this case—while the other variables remain fixed, and the whole cycle is repeated until the algorithm converges. The alternating variable method applied to this case could be efficient, because the two variables (gain and phase) are completely decoupled and the change of objective function occurs through the principal axes of the ellipse. However, the measured out-of-band power at each iteration is subject to the noise, since an estimation is used with a finite number of symbol integrations. Therefore, the simplex algorithm, which is known to be robust when the objective function suffers from substantial noise, is suggested [11]. A simplex is a set of $N+1$ equidistant points in N -dimension space R^N —an equilateral triangle in this case. An initial simplex is chosen at the beginning of the iteration. The vertex at which the objective function value is largest is determined, and a new simplex is formed by mirroring this vertex through the other two, as illustrated in Figure 3.8. The process is repeated and eventually no further progress is observed, owing to the insufficient resolution of the simplex.

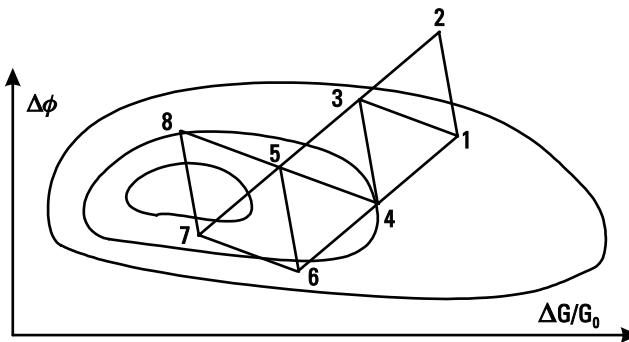


Figure 3.8 Illustration of the simplex method with two variables [11].

Then a smaller simplex is constructed, allowing the iteration to continue until the objective function value is under an acceptably low level. In this method, the correction of imbalance relies on the measurement of the low-power out-of-band emission, which demands a long integration time for each adaption. This requirement sets a lower limit on the calibration time of approximate 1–2 seconds, which is a consideration in real-time applications.

3.3.3 Direct Search Algorithm

The performance of the simplex algorithm may be degraded if the phase imbalance is the combined result of the phase imbalance due to differences in electrical length and the gain imbalance through AM-PM conversion of the nonlinear power amplifier. A direct search method was proposed in [12] to solely correct the gain imbalance as well as the consequent phase imbalance. The block diagram for the baseband model of this approach is shown in Figure 3.9. The amplifier outputs are combined, attenuated, and then fed to a LPF to suppress the out-of-band spectrum. The attenuated output signal can be expressed as

$$\begin{aligned} S(t) &= \frac{1}{2} [s(t) - e(t)] + \frac{1}{2} [1 + g \cdot e^{j\theta(g)}] [s(t) + e(t)] \\ &= s(t) + \frac{1}{2} g e^{j\theta(g)} [s(t) + e(t)] \end{aligned} \quad (3.33)$$

where g is the gain imbalance and $\theta(g)$ is the AM-PM induced phase imbalance. Referring to Figure 2.21, it is clear that for most linearly modulated signals, especially for high-order modulations such as 16-QAM,

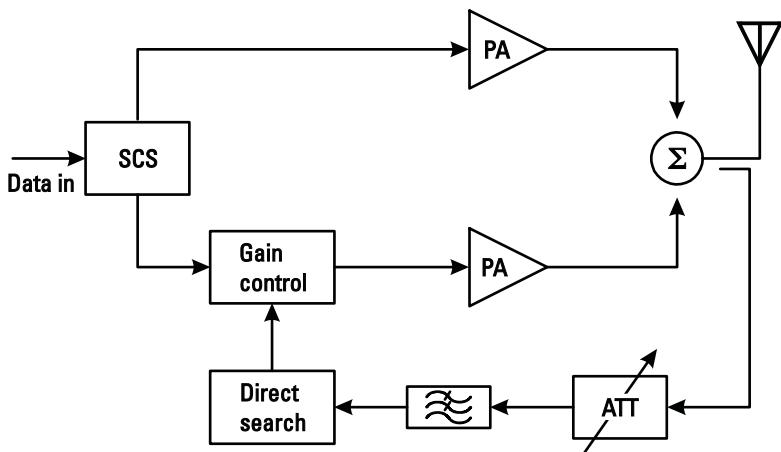


Figure 3.9 Baseband block diagram for direct search algorithm [12].

the quadrature signal $e(t)$ dominates the inband power other than the source signal $s(t)$. The lowpass filtering suppresses the out-of-band power but hardly affects the inband distorted signal power due to $e(t)$. A reference modulated source signal $s_r(t)$ is then subtracted from the filter output [12],

$$S_{\text{sub}}(t) = \frac{1}{2} g e^{j\theta(g)} [s(t) + e(t)] \quad (3.34)$$

Note that $s(t) + e(t)$ has constant amplitude r_{\max} ; thus, this result can be used by a direct search algorithm to find the gain imbalance for correction in the lower amplifier branch. This technique is based on the evaluation of the inband distortion by downconverting the combined output and subtracting it from the input signal with an extra D/A branch. The subtraction has to be precise to obtain the complete cancellation of the source signal. Moreover, no phase imbalance due to electrical length difference between two amplifier branches is assumed.

3.3.4 Background Calibration Algorithm

The background calibration algorithm is an improved feedback approach over the foreground one, which operates continuously in background during regular data transmission. Instead of generating a set of calibration signals, the background scheme makes use of the transmitted data as a kind of calibration signal. As shown in Figure 3.10, a small portion of the amplifier output is coupled into the feedback loop and downconverted by a mixer.

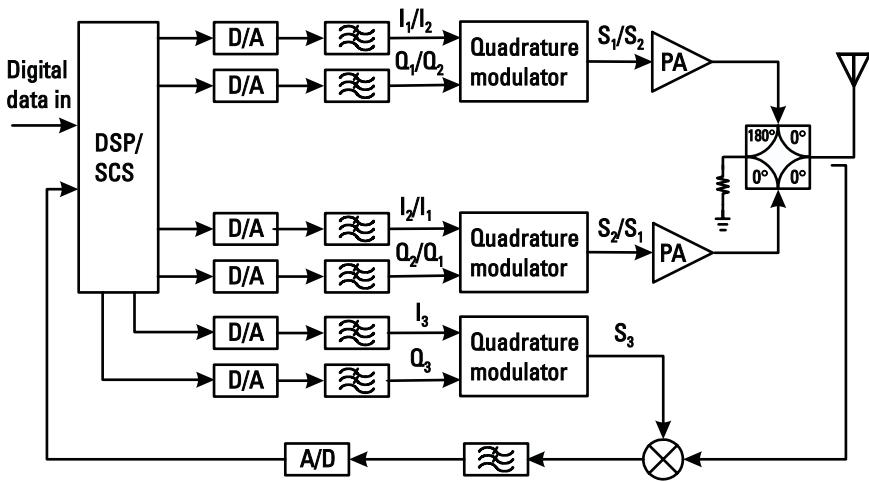


Figure 3.10 Block diagram for background calibration algorithm [9].

Then the mixed signal is lowpass filtered, A/D converted, and sent back to the DSP/SCS. The baseband digital circuit controls the third branch signal such that S_3 connects to either of the upper or lower amplifier branch inside the DSP. Furthermore, in each case the DSP regularly exchanges two component vectors S_1 and S_2 in two amplifier branches back and forth. The gain and phase imbalance are completely characterized by determining the maximum and minimum signal values.

3.3.4.1 Theory of Background Calibration Algorithm

Assume that G_0 is the amplifier gain and ϕ_0 is the path delay and that the gain and phase imbalance are $\Delta G/G_0$ and $\Delta\phi$, respectively. Then before and after the DSP exchanges S_1/S_2 , the combined output may be expressed as

$$S(t) = G_0 V_m \cos \left[\omega_c t + \theta(t) \mp \psi(t) + \phi_0 \right] + \left(1 + \frac{\Delta G}{G_0} \right) G_0 V_m \cos \left[\omega_c t \pm \theta(t) + \psi(t) + \phi_0 + \Delta\phi \right] \quad (3.35)$$

where “ \mp ” and “ \pm ” are the consequences of exchanging two component vectors S_1 and S_2 . Note that as long as the path imbalance is maintained under a certain low level, the exchange between S_1 and S_2 makes little difference to the amplifier output signal. When the S_3 branch connects to the upper branch inside the DSP, called state “A,” we have

$$S_{3A}(t) = V_3 \cos [\omega_c t + \theta(t) \mp \psi(t) + \phi_3] \quad (3.36)$$

where V_3 stands for the signal amplitude of the mixer branch, and ϕ_3 is the phase delay. The downconversion mixer multiplies the outphasing amplifier output (3.35) with (3.36), and after lowpass filtering we obtain

$$S_{\text{LPF}} = \frac{1}{2} G_L V_m \cos \phi_L + \frac{1}{2} \left(1 + \frac{\Delta G}{G_0}\right) G_L V_m \cos [\pm 2\psi(t) + \phi_L + \Delta\phi] \quad (3.37)$$

where G_L is the loop gain and ϕ_L consists of the loop phase delay and the phase shift introduced by the downconversion mixer. Recall that $\cos[\psi(t)] = a(t)/V_m$, which is the only time variable in (3.37); hence the detected signal is a sinusoid modulated by baseband amplitude $r(t)$ and offset by a dc constant. The dc offset and the amplitude of this sinusoidal function carry the gain and phase imbalance, and they can be extracted by determining the two extreme cases—the maximum and minimum signal values. This function is accomplished by the baseband DSP. The band-limited characteristics of the baseband input implies that the amplitude $a(t)$ has a large variation with time. As stated previously, $0^\circ \leq \psi(t) \leq 90^\circ$; therefore we have

$$-180^\circ \leq \pm 2\psi(t) \leq 180^\circ \quad (3.38)$$

The above expression guarantees that the maximum and minimum signal values can be found. The exchanging of S_1 and S_2 may create an overshoot on the output envelope input to the DSP, due to the phase discontinuity in the S_3 branch before and after exchanging, which in turn degrades the measurement accuracy of the maxima and minima. This error effect can be minimized by exchanging S_1 and S_2 when these two vectors are close to each other [i.e., $a(t) \approx V_m$], or possibly hold the DSP/SCS output at a constant for a short period of time until the LPF reaches the steady state. Now, two quantities are obtained by combining these maxima and minima—that is,

$$S_{A+} = \max_A + \min_A = G_L V_m \cos \phi_L \quad (3.39)$$

$$S_{A-} = \max_A - \min_A = \left(1 + \frac{\Delta G}{G_0}\right) G_L V_m \quad (3.40)$$

Similar procedures applied to the state “B,” when the DSP switches S_3 to the lower amplifier branch. Specifically, S_3 connects to the lower branch, and when mixed with the outphasing amplifier output, we obtain

$$S_{LPF} = \frac{1}{2} \left(1 + \frac{\Delta G}{G_0} \right) G_L V_m \cos(\phi_L + \Delta\phi) + \frac{1}{2} G_L V_m \cos \left[\mp 2\psi(t) + \phi_L \right] \quad (3.41)$$

Similarly, S_{B+} and S_{B-} can be computed,

$$S_{B+} = \left(1 + \frac{\Delta G}{G_0} \right) G_L V_m \cos(\phi_L + \Delta\phi) \quad (3.42)$$

$$S_{B-} = G_L V_m \quad (3.43)$$

Comparing (3.40) and (3.43), we immediately obtain the gain imbalance

$$\frac{\Delta G}{G_0} = \frac{S_{A-}}{S_{B-}} - 1 \quad (3.44)$$

The determination of phase imbalance is a little bit more involved, since this quantity is resolved from $\cos(\phi_L + \Delta\phi)$ and $\cos \phi_L$, which are calculated by the other two ratios of (3.40) and (3.43); that is,

$$\Delta\phi \approx \frac{\cos \phi_L - \cos(\phi_L + \Delta\phi)}{\sin \phi_L} \quad (3.45)$$

The resolution of $\Delta\phi$ from $\cos(\phi_L + \Delta\phi)$ strongly depends on ϕ_L since ϕ_L is a small quantity. This situation occurs due to the finite word length representation of the voltage waveform by the ADC, and hence limits the SNR in the calibration loop. The optimum resolution occurs as $\sin \phi_L = \pm 1$. This condition can be achieved by monitoring the ratio $\cos \phi_L = S_{a+}/S_{b-}$ and introducing a proper phase shift in S_3 branch such that $S_{a+}/S_{b-} \approx 0$. As a fairly rough estimation, we have

$$\Delta\phi \approx \pm \left(\frac{S_{B+}}{S_{A-}} - \frac{S_{A+}}{S_{B-}} \right) \quad (3.46)$$

where “ \pm ” is determined by the sign of ϕ_L . In a similar manner to P_L in the foreground calibration case, $\sin \phi_L$ only scales the measured phase imbalance. Hence the accurate determination of $\sin \phi_L$ is not critical and is unnecessary, and (3.46) is effective in an iterative sense, even if ϕ_L cannot be well-controlled.

3.3.4.2 Practical Considerations and Algorithm Limitations

The quadrature modulator error is more complicated than in the foreground case, since there are three I/Q modulator branches. For the sake of brevity, we will assume that all the quadrature errors are equal [i.e., $g_{1,2,3} = \delta_{1,2,3} = \delta$]. This will simplify the calculations; yet the final results will provide general directions. The detailed derivations are neglected here, and the worst case of the four measured signals is given by

$$S_{A+} = (1 + \delta)G_L V_m \cos \phi_L \quad (3.47)$$

$$S_{A-} = (1 + \delta)(1 + \frac{\Delta G}{G_0})G_L V_m + 2\sqrt{2}\delta G_L V_m \cos \phi_L \quad (3.48)$$

$$S_{B+} = (1 + \frac{\Delta G}{G_0})(1 + \delta)G_L V_m \cos(\phi_L + \Delta\phi) \quad (3.49)$$

$$S_{B-} = (1 + \delta)G_L V_m + 2\sqrt{2}\delta G_L V_m \cos \phi_L \quad (3.50)$$

The gain imbalance is calculated according to (3.44)

$$\begin{aligned} \left(\frac{\Delta G}{G_0} \right)_{\text{est}} &= \frac{S_{A-}}{S_{B-}} - 1 \\ &\approx \left(\frac{\Delta G}{G_0} \right)_{\text{act}} \end{aligned} \quad (3.51)$$

and the phase imbalance can be calculated by

$$\begin{aligned} (\Delta\phi)_{\text{est}} &= \pm \left(\frac{S_{B+}}{S_{A-}} - \frac{S_{A+}}{S_{B-}} \right) \\ &\approx (\Delta\phi)_{\text{act}} \left(1 - 2\sqrt{2}\delta \cos \phi_L \right) \end{aligned} \quad (3.52)$$

From this very rough approximation, the quadrature errors of the I/Q modulators will not affect the measurement of the gain imbalance, while the measurement error of the phase imbalance has the same order as the quadrature errors. As we know, $\cos \phi_L$ is kept close to zero for the optimum estimation on the phase imbalance and that reduces the effects of quadrature errors.

The characterization of the gain and phase imbalance takes advantage of the time-varying characteristics of the baseband input signal. In other words, if the baseband input is a constant amplitude, this calibration scheme fails. However, there is no reason to use an outphasing amplifier to amplify such

a constant amplitude signal, so this is not a limitation in practice. Referring to (3.37), in the case of optimum condition, to guarantee that the DSP finds the maxima and minima, the amplitude variation has to satisfy the following expression

$$\frac{\max[a(t)]}{\min[a(t)]} > \sqrt{2} \quad (3.53)$$

This situation is illustrated in Figure 3.11, in which at least a $\pm 90^\circ$ variation on $\pm 2\psi(t)$ is required in order that the function passes through its maxima and minima from the optimum point. In case the S_3 branch is not well-controlled, the variation of the signal amplitude has to increase accordingly to accommodate the deviation from the optimum point. Equation (3.53) is thus the minimum requirement and is applicable to most practical applications.

The dc offset of the mixer, the LPF, and the ADC add to the measured signal values as a constant background and degrade the measurement accuracy. Note, however, that the dc offset will not affect the accuracy of the measured gain imbalance, since it cancels out in S_{A-} and S_{B-} . To take it into account, this value is determined and subtracted from S_{A+} and S_{B+} . The dc offset may be simply measured by disabling the transmitted signal.

The quantization error of the A/D converter degrades the measurement accuracy and perhaps the algorithm convergence. The bit length of the A/D converter must be minimized to reduce the computation load of the DSP. The worst-case estimation errors of the gain and phase imbalance are bounded by

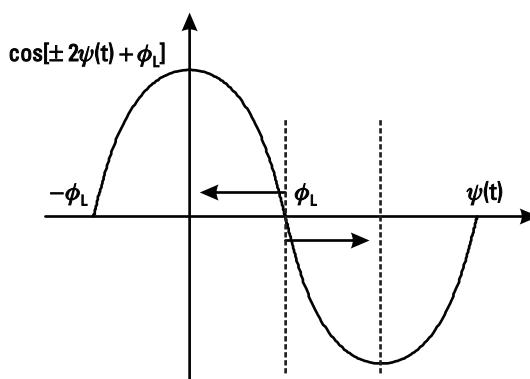


Figure 3.11 Minimum requirement on the amplitude variation of the baseband signal.

$$\frac{\Delta G}{G_0 \text{ est}} \approx \frac{\Delta G}{G_0 \text{ act}} \pm \frac{2\Delta}{G_L V_m} \quad (3.54)$$

$$\Delta\phi_{\text{est}} \approx \Delta\phi_{\text{act}} \pm \frac{4\Delta}{G_L V_m} \quad (3.55)$$

As it turns out, the word length of the A/D converter must satisfy

$$B + 1 = 2 - \frac{\log_{10} \delta}{\log_{10} 2} \quad (3.56)$$

where δ is the maximum allowable gain or phase imbalance. For the allowable imbalance better than 2° for phase or 0.3 dB for gain, a 7-bit word length representation would be adequate, which corresponds to around -40 dBc ACI for CDMA IS-95.

The background calibration is transparent to the data transmission; hence its application is not limited by the communication standard. However, compared to the foreground calibration, the standard background scheme requires an extra I/Q modulator branch to downconvert the system output and extract the imbalance information, including two ADCs, two reconstruction filters, and one I/Q modulator. This added complexity becomes even more complicated when considering the matching among these three branches. By carefully examining the background algorithm, it is clear that only two different RF branch signal components are taken and mixed with the amplifier output to calibrate the system— S_1 and S_2 . This suggests that the extra I/Q modulator branch may be replaced by two RF switches—an alternative implementation approach of the background calibration scheme, as illustrated in Figure 3.12. There are various

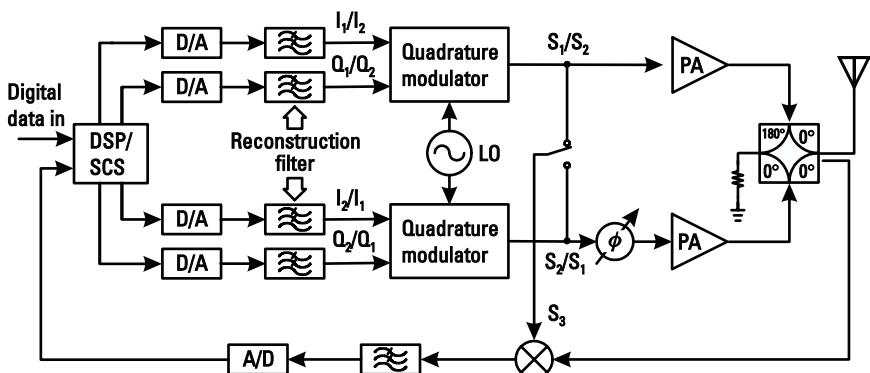


Figure 3.12 Background calibration implemented with RF switches [9].

commercially available RF diode switches. Alternatively, the RF switch may be implemented with MEMS technology. Note that the compensation of the phase imbalance is accomplished by the phase shifter instead of predistortion inside the DSP. A phase shift in baseband results in no change of the measured phase imbalance, even if it has been predistorted and compensated.

Since the RF switch couples two component vectors to the down-conversion mixer, any mismatch between the state “A” and “B” in the S_3 branch may be directly transformed to the mixer output and create measurement error. Since the mixer can operate in saturation, its output is not sensitive to the relatively small amplitude variation of the LO driving signal. Moreover, though the RF switch branch requires careful phase matching, detailed analysis shows that the phase difference between the two switch states affects only the measured phase imbalance, not the gain imbalance, and the measured phase imbalance is given by

$$\Delta\phi_{\text{est}} = \Delta\phi_{\text{act}} - \delta \quad (3.57)$$

Hence, the final phase imbalance converges to δ —the phase mismatch between the state “A” and “B.”

Figure 3.13 displays a snapshot of the lowpass-filtered signal during the state “B.” The abrupt transition in the middle of the graph corresponds to the moment of exchanging between S_1 and S_2 . The quantity ϕ_L in this case was kept around 70° . Figure 3.14 shows the simulated waveform after lowpass filtering with $\phi_L = 70^\circ$. A short period of data ($20 \mu\text{s}$) is also displayed as an inset in Figure 3.13 to illustrate the rapid variation of the lowpass filtered signal.

An example of the improvement is shown in Figure 3.15. Without correction, the ACI is around -28 dB , while with correction the out-of-band spectrum is suppressed below -35 dBc . The measured gain and phase imbalance prior to calibration are 0.6 dB and 8° , respectively.

3.4 Mismatch Correction Scheme for Broadband Applications

Simple correction of the gain and phase imbalance as described in previous sections is no longer effective in broadband applications such as WCDMA, orthogonal frequency division multiplexing (OFDM), and multicarriers in base stations, since a high degree of matching between two amplifier branches are demanded throughout the entire channel. The equalization technique may be employed to compensate the mismatch of the two branches

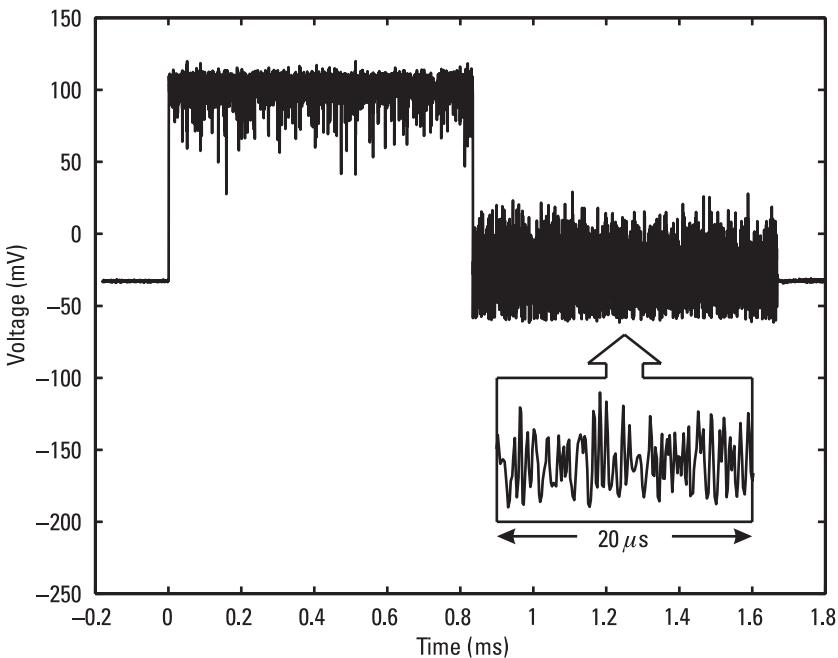


Figure 3.13 Measured lowpass-filtered signal during calibration [9].

and balance their frequency responses to the desirable characteristics. This can be achieved by incorporating equalization in two amplifier channels prior to the power amplification and adaptively adjusting the equalizer coefficients for the minimum the signal distortion [13, 14]. In a typical implementation of this generalized correction scheme, the calibration and channel equalization are performed digitally inside the baseband DSP circuit, with the distorted output to guide the algorithm adaption.

The calibration algorithm consists of two iterative operations—system identification and adaptive equalization. The power amplifier transfer characteristic is needed for channel equalization, and this is accomplished by system identification. During system identification, a digital model of the amplifier system is constructed, and the output signal of the actual system is A/D converted and compared to the model output. Then the error signal $e[n]$ is taken by the least square (LS) algorithm [15] as the cost function to find the transfer characteristics of two power amplifiers, as shown in Figure 3.16. Due to the near constant envelope feature of the amplifier driving signals, two nonlinear power amplifiers are conveniently modeled as finite impulse response (FIR) filters. The filter coefficients are extracted by minimizing the cost function in a least-square sense; that is,

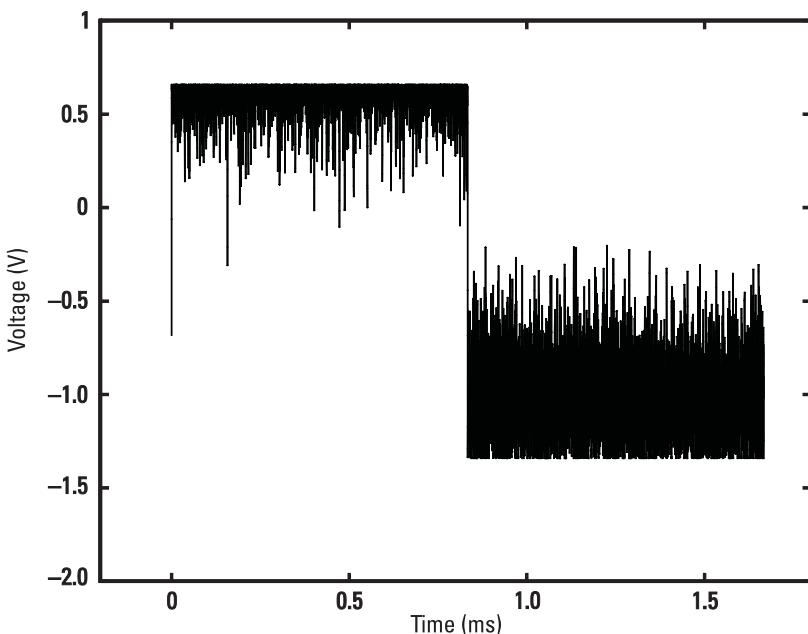


Figure 3.14 Simulated lowpass-filtered signal during calibration ($\phi_L = 70^\circ$).

$$\varepsilon_1[n] = |s_{\text{out}}[n] - \hat{s}_{\text{out}}[n]|^2 \quad (3.58)$$

Figure 3.17 illustrates the adaptive equalization process for computing the two equalizer coefficients. This is done by interchanging the power amplifier model and the equalizer, comparing the baseband input to the system output and minimizing the cost function

$$\varepsilon_2[n] = |s_{\text{in}}[n] - \hat{s}_{\text{out}}[n]|^2 \quad (3.59)$$

The amplifier input and equalizer output can also be taken into account as a part of the weighted cost function. Readers may refer to [13, 14] for detailed discussion. The algorithm shows significant improvement on the system linearity for four CDMA carriers [14], though intensive DSP and adaptive filtering techniques are involved. The 16–24-tap FIR filters provide sufficient bandwidth and null depth for practical implementation.

3.5 VCO-Derived Synthesis

PLL technology has been extensively used in many digital and analog applications, for example, frequency synthesis, modulation, demodulation,

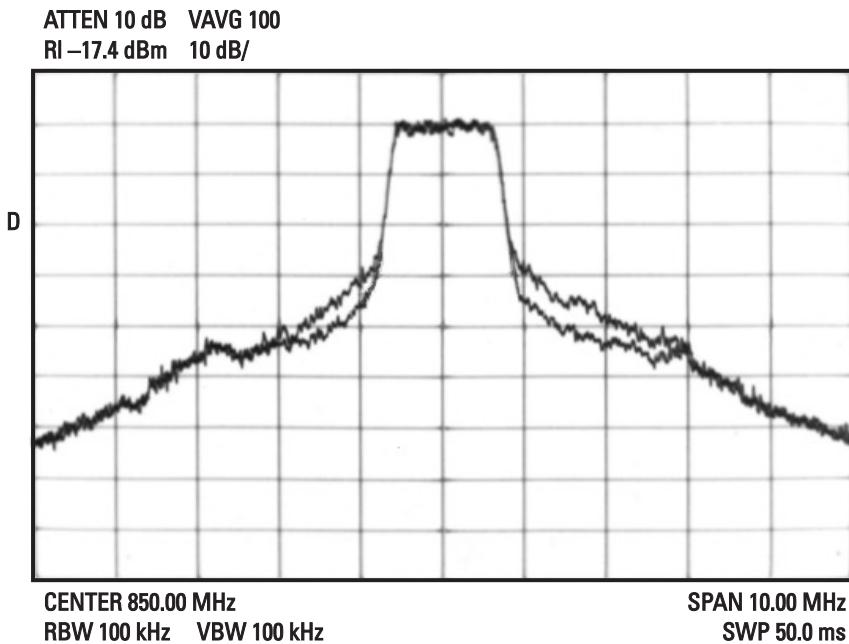


Figure 3.15 Measured output spectra for CDMA IS-95 with and without calibration using background correction [9].

and clock recovery. A basic PLL comprises a phase detector, a loop filter, and a VCO, as shown in Figure 3.18. The phase detector compares the input reference and the VCO output, the phase error is developed and then fed to the VCO through the loop filter. The negative feedback of the control system drives the VCO in such a manner that the phase error is eventually reduced. Different types of loop filters can be incorporated into the system depending on the applications to improve the loop static and dynamic performance. Once the loop lock is established, the VCO output is said to be “phase-locked” to the input reference, with a fixed phase relationship (usually 0° or 90° phase offset) determined by the nature of the phase detector and loop filter.

The original VLL, as an extended concept of the feedback signal generator with both phase and magnitude locked, was proposed in [16]. That particular circuitry, though capable of simultaneously tracking the phase and magnitude of the reference signal, has certain limitations and has failed to gain a widespread adoption. Two more versatile topologies—an improved VLL [17] and a similar approach called CALLUM [18–23]—were developed afterward to overcome the shortcomings of the primitive version with expanded applications. These topologies utilize two cross-coupled PLLs, in

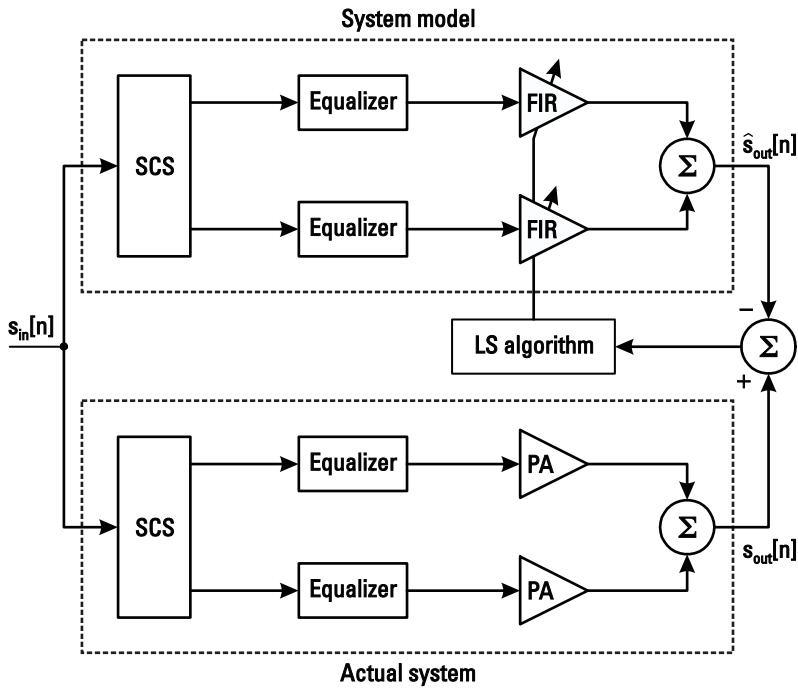


Figure 3.16 System identification to estimate the amplifier characteristics [14].

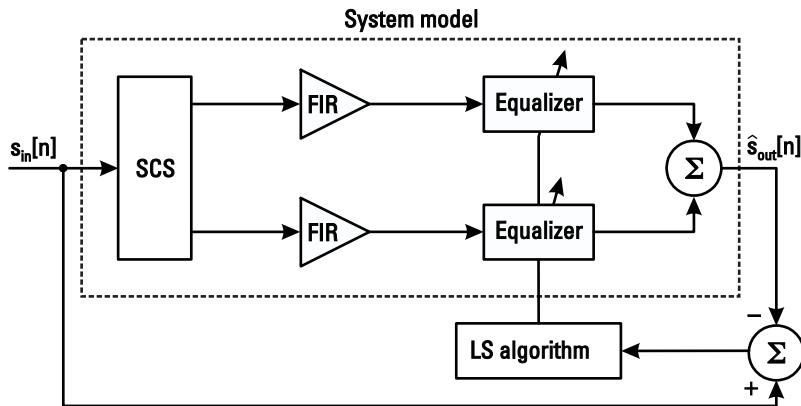


Figure 3.17 Adaptive equalization to compute equalizer coefficients [14].

which each VCO is driven by one of two error-control signals and two VCOs' output are combined, as illustrated in Figure 3.19. The error-control signals are the function of the absolute phase and magnitude errors between

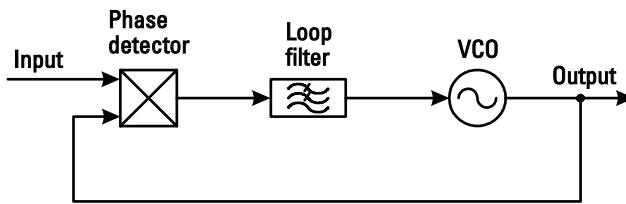


Figure 3.18 Basic PLL block diagram.

input reference and output signal in the case of VLL, or the absolute errors of the real and imaginary parts of the input reference and output signal in the case of CALLUM. With this cross-coupled configuration, an arbitrary input varying in phase and amplitude can be synthesized. Note that due to the constant-amplitude nature of the VCO's output, this circuit topology is analogous to the outphasing approach, except that the signal component separation is realized by the feedback-controlled signal generator. By adding the power gain stages following the VCOs, the locked loop naturally turns into a linear power amplifier. In this respect, the VLL and CALLUM can be regarded as particular implementations of the outphasing amplifier system. A unique advantage of CALLUM and VLL, over the conventional outphasing amplifier, is the capability to compensate for the phase and amplitude imbalance between the two VCO branches. This benefit, which naturally results from the loop negative feedback, eases the matching issue and is crucial for the successful practical implementation of the VCO-derived approach. However, due to the closed-loop high-gain nature, particularly the excess phase shift and time delay introduced in the RF signal path, CALLUM and VLL are inherently narrowband. Their applications, therefore, concentrate on low-cost and power-efficient integrated circuit implementations for portable equipment.

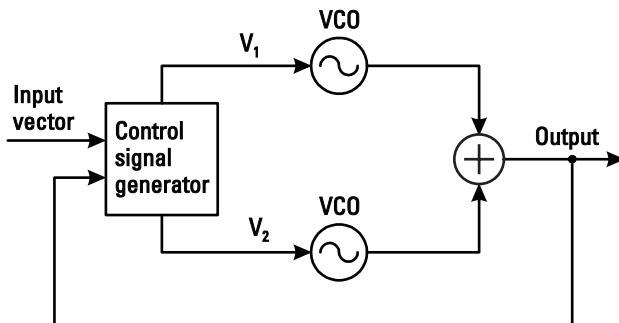


Figure 3.19 Block diagram of generalized locked loop.

To derive the expressions of the VCO control voltages in terms of the reference input and error signals, consider a general complex baseband input signal,

$$s(t) = r(t)e^{j\theta(t)} \quad 0 \leq r(t) \leq r_{\max} \quad (3.60)$$

The two VCOs' output signals have the phase in identical form to that of an outphasing system; that is,

$$\theta_1(t) = \theta(t) - \psi(t) \quad (3.61)$$

$$\theta_2(t) = \theta(t) + \psi(t) \quad (3.62)$$

where

$$\psi(t) = \cos^{-1}\left[\frac{r(t)}{r_{\max}}\right] \quad (3.63)$$

The VCO free-running frequency may be ignored here, in the case that frequency translation is involved—for simplicity without losing the generality of discussion. Taking the derivative with respect to time on both sides of (3.61) and (3.62), we have

$$\theta'_1(t) = \theta'(t) + \frac{r'(t)}{\sqrt{r_{\max}^2 - r^2(t)}} \quad (3.64)$$

$$\theta'_2(t) = \theta'(t) - \frac{r'(t)}{\sqrt{r_{\max}^2 - r^2(t)}} \quad (3.65)$$

The VCO can be treated as an integrator, and the VCO output phase $\theta_{1,2}$ is related to the input driving voltage $V_{1,2}$ by

$$\theta'_{1,2}(t) = c \cdot V_{1,2}(t) \quad (3.66)$$

where c is the VCO sensitivity. For a circuit implementation, the quantities $\theta'(t)$ and $r'(t)$ are difficult to implement. Note that $\theta'(t)$ and $r'(t)$ are proportional to $d\theta$ and dr and may be approximated by the phase and magnitude tracking errors— $\Delta\theta$ and Δr . This corresponds to the situation in which the loop is in lock, while the changes of phase and amplitude of the input reference vector are small [20]. Thus, the above equations relate

the VCO control voltages to the phase and magnitude errors. Equations (3.64) and (3.65) are therefore the basis of VLL, and the controlled feedback loop is implemented in polar form.

Alternatively, the feedback locked loop can be implemented in Cartesian form as in the CALLUM case, by expressing signal amplitude and phase— $r(t)$ and $\theta(t)$ —in terms of real and imaginary parts— $x(t)$ and $y(t)$,

$$r(t) = \sqrt{x^2(t) + y^2(t)} \quad (3.67)$$

$$\theta(t) = \tan^{-1} \frac{y(t)}{x(t)} \quad (3.68)$$

Substitute (3.67) and (3.68) into (3.64) and (3.65), and (3.69) and (3.70) can be derived:

$$\theta'_1(t) = -\left(\frac{y}{r^2} - \frac{x}{r\sqrt{r_{\max}^2 - r^2}}\right)x' + \left(\frac{x}{r^2} + \frac{y}{r\sqrt{r_{\max}^2 - r^2}}\right)y' \quad (3.69)$$

$$\theta'_2(t) = -\left(\frac{y}{r^2} + \frac{x}{r\sqrt{r_{\max}^2 - r^2}}\right)x' + \left(\frac{x}{r^2} - \frac{y}{r\sqrt{r_{\max}^2 - r^2}}\right)y' \quad (3.70)$$

In a similar manner to a phase-locked loop (PLL), as long as the loop lock is maintained, quantities $x'(t)$ and $y'(t)$ are proportional to dx and dy and can be approximated by the absolute errors Δx and Δy . Equations (3.69) and (3.70) are thus the basis of CALLUM.

3.5.1 CALLUM

The implementation of (3.69) and (3.70) results in a generic form of CALLUM, so-called CALLUM1. The system control equations are described by [20]

$$\frac{d\theta_1}{dt} = -\frac{k}{c} \left(\frac{y}{r^2} - \frac{x}{r\sqrt{r_{\max}^2 - r^2}} \right) \Delta x + \frac{k}{c} \left(\frac{x}{r^2} + \frac{y}{r\sqrt{r_{\max}^2 - r^2}} \right) \Delta y \quad (3.71)$$

$$\frac{d\theta_2}{dt} = -\frac{k}{c} \left(\frac{y}{r^2} + \frac{x}{r\sqrt{r_{\max}^2 - r^2}} \right) \Delta x + \frac{k}{c} \left(\frac{x}{r^2} - \frac{y}{r\sqrt{r_{\max}^2 - r^2}} \right) \Delta y \quad (3.72)$$

where k is the loop gain. Figure 3.20 illustrates the block diagram of CALLUM1. The control signal generator scales the absolute error signals and combines them to generate two VCO control voltages according to the above control equations. Equations (3.71) and (3.72) are first-order nonlinear differential equations. In general, such equations are difficult to solve analytically, and numerical methods have to be used. When the system is in lock, the nonlinear system can be described by a linearized model for small changes in the phase and amplitude of the input vector. This is exactly the same situation as in a PLL. In the case of CALLUM1, the linearized model is simply a single-pole system with a time constant of $1/k$. The phase and amplitude responses have been analyzed in [20] by assuming they are uncorrelated. One important result is the maximum locking frequency when the system is driven by a sinusoidal signal with constant amplitude, given by

$$f_{\max} = \pm \frac{k}{2\pi r} \sqrt{r_{\max}^2 - r^2} \quad (3.73)$$

The dynamic behavior of CALLUM1 was briefly studied in [23], which shows acquisition process is highly dependent on the initial conditions. A small perturbation in the initial conditions could result in a significant difference in acquisition behavior.

The automatic compensation mechanism is obvious from the negative feedback that minimizes the phase and amplitude errors and duplicates

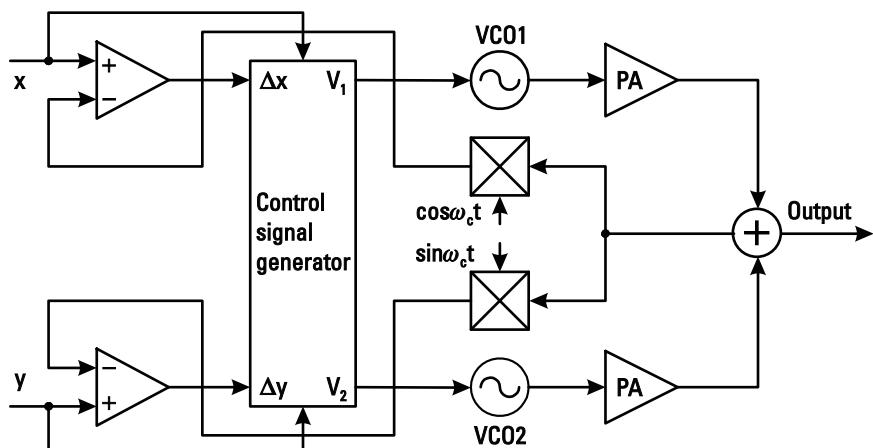


Figure 3.20 Block diagram of CALLUM1 [20].

the synthesized output to the input reference. Detailed analysis indicates nonzero phase and amplitude errors in the general cases. The tracking performance may be improved by enclosing two integrators in the error signal paths. This is a similar situation as in a PLL.

The accurate implementation of the control signal generator in CALLUM1 is a task of equal difficulty to the SCS itself in an outphasing system. Several variations of CALLUM have been proposed to reduce the implementation complexity for control signal generation. In the first modification, the denominators in (3.71) and (3.72) are completely ignored; this simplified version is referred to as CALLUM2. The system control equations for CALLUM2 are thus described by

$$\frac{d\theta_1}{dt} = \frac{k}{c}(x - y)\Delta x + \frac{k}{c}(x + y)\Delta y \quad (3.74)$$

$$\frac{d\theta_2}{dt} = -\frac{k}{c}(x + y)\Delta x + \frac{k}{c}(x - y)\Delta y \quad (3.75)$$

and the block diagram is shown in Figure 3.21. Eliminating the divisions in CALLUM1 reduces the complexity and allows CALLUM2 to be implemented in low-cost and power-efficient analog circuitry. However, the sensitivity and tracking performance are degraded, since now the overall system loop gain varies with the input signal amplitude. This is apparent by comparing the control equations of CALLUM1 and CALLUM2 in the case of a low-input envelope. As the signal amplitude decreases, the tracking performance continuously degrades, and eventually the loop may fail to lock. At a high-signal amplitude, on the other hand, the increased gain may cause

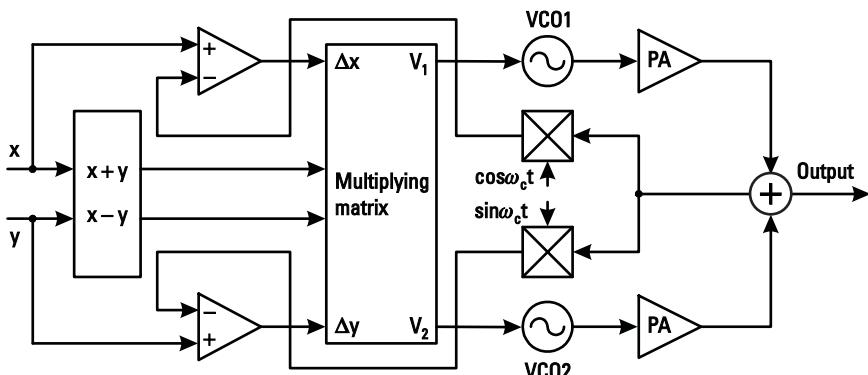


Figure 3.21 Block diagram of CALLUM2 [20].

the loop to become unstable. The maximum frequency of CALLUM2 for a constant amplitude-sinusoidal input is given by [20]

$$f_{\max} = \pm \frac{k}{2\pi} r \sqrt{r_{\max}^2 - r^2} \quad (3.76)$$

The basic CALLUM system, proposed in [18], utilizes two VCOs driven by two separate Cartesian component error signals, as shown in Figure 3.22, reminding us of the classical Cartesian feedback power amplifier. This architecture, though simple in appearance, suffers from the fact that the stability region covers only half of the complex plane. Specifically, the system is stable only if [19]

$$-\frac{3\pi}{4} \leq \tan^{-1} \frac{y}{x} \leq \frac{\pi}{4} \quad (3.77)$$

Look back at the control equations of CALLUM1, and it can be seen that the basic CALLUM utilizes an incomplete set of signals to drive the VCOs. Sign-control switching is necessary to ensure stable operation for arbitrary phases as well as a smooth transition. This improved system, called CALLUM3, is shown in Figure 3.23. The control equations are given by [20]

$$\frac{d\theta_1}{dt} = -\frac{k}{c} \text{sign}[y] \Delta x \quad (3.78)$$

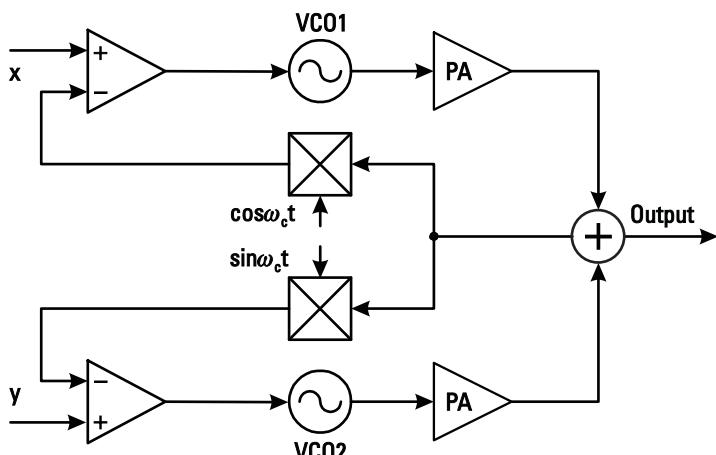


Figure 3.22 Basic structure of CALLUM [18].

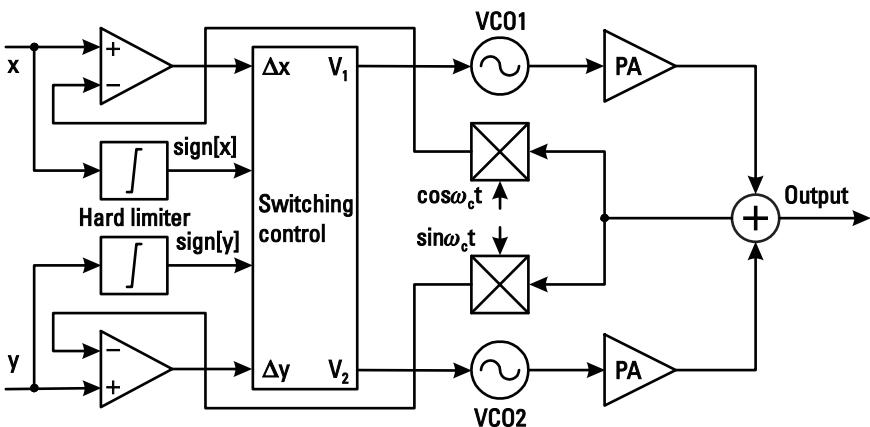


Figure 3.23 Block diagram of CALLUM3 [20].

$$\frac{d\theta_2}{dt} = \frac{k}{c} \text{sign}[x]\Delta y \quad (3.79)$$

CALLUM4 is a further improvement on CALLUM3. In CALLUM3, the phase response time constant is optimum when the input phase is $45^\circ + n \times 90^\circ$. CALLUM4 thus encloses a phase shifter to maintain the phase of the absolute error signals to always be 45° . This is accomplished by a phase shifter that negatively mirrors the phase of the input signal. In this way, the system remains simple and always achieves optimum operation while eliminating the need for switching. Figure 3.24 illustrated the block diagram of CALLUM4.

A major disadvantage of CALLUM is its sensitivity to the input signal amplitude. The loop fails to remain locked when the input signal amplitude drops to zero, as in a classical two-tone test and a high-level modulated signal. As the input amplitude goes through zero, loop lock is temporarily lost and the system undergoes a reacquisition process. Moreover, when the input signal amplitude remains zero, the two VCOs are free-running and uncorrelated, which means that their phase relation is arbitrary; therefore the summed output signal amplitude and phase are undetermined. Unless the two amplifier paths are perfectly matched, there is no way to generate zero output signal. This is a common nature of all RF synthesis approaches including VLL. One way to solve this is to bias the input vector such that it never falls to zero and then remove the offset at the output. This could be done by use of a three-VCO structure [21]—however, at the cost of increased complexity.

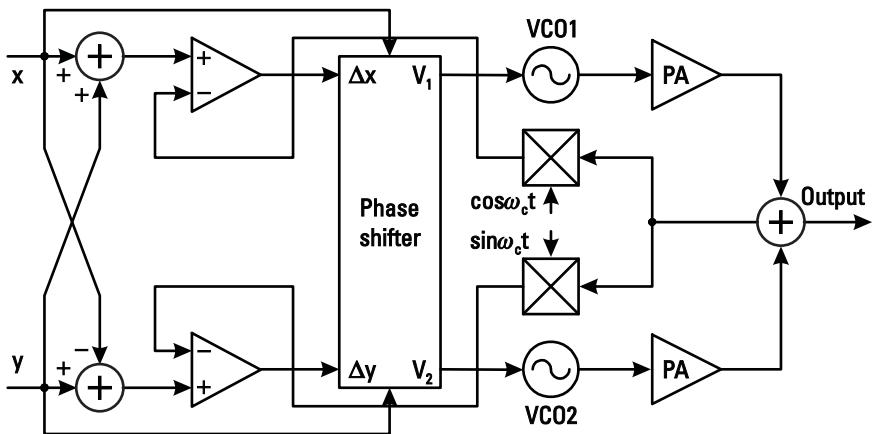


Figure 3.24 Block diagram of CALLUM4 [20].

3.5.2 VLL

From (3.64) and (3.65), the generic control equations for a VLL may be described as

$$\frac{d\theta_1}{dt} = k_\theta \Delta\phi + \frac{k_r \Delta r}{\sqrt{r_{\max}^2 - r^2}} \quad (3.80)$$

$$\frac{d\theta_2}{dt} = k_\theta \Delta\phi - \frac{k_r \Delta r}{\sqrt{r_{\max}^2 - r^2}} \quad (3.81)$$

where k_θ and k_r are the loop gains, and they are not necessarily the same. The implementation in [17] of the VLL is a simplified version in which the divisions in the control equations are eliminated; that is,

$$\frac{d\theta_1}{dt} = k_\theta \Delta\phi + k_r \Delta r \quad (3.82)$$

$$\frac{d\theta_2}{dt} = k_\theta \Delta\phi - k_r \Delta r \quad (3.83)$$

and this approach is illustrated in Figure 3.25. The VCO's outputs are power-amplified and combined to synthesize the desired phase and amplitude characteristics of the input reference. A directional coupler samples the final output and feeds it back to a phase detector and a magnitude detector. The detectors generate the phase and magnitude difference as

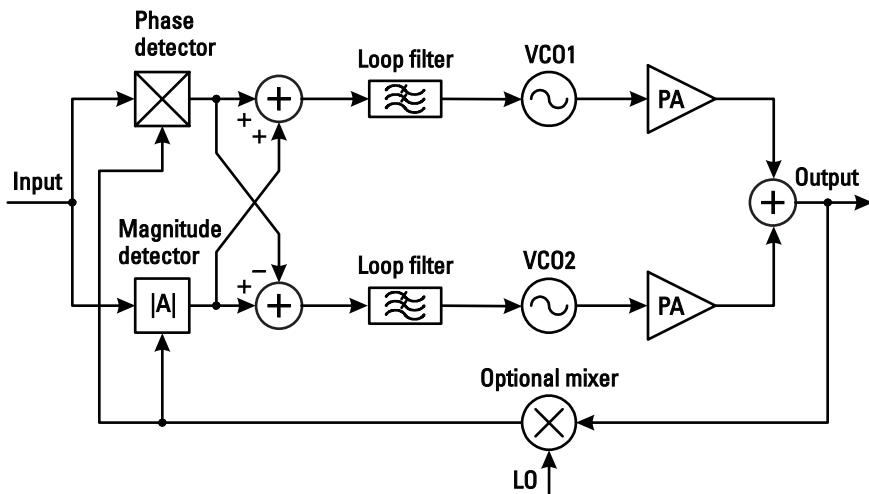


Figure 3.25 VLL [17].

the error signals. The two error signals are then summed and subtracted to constitute two driving signals for the VCOs through the loop filters. When the loop is maintained in lock, the phase and magnitude errors are minimized, and the system output tracks the phase and amplitude of the input reference signal. A simple downconversion mixer can be placed in the feedback loop for frequency translation, and the carrier frequency is set by the downconverter LO.

The control equations may be simplified by rearranging (3.82) and (3.83) as follows:

$$\frac{d(\theta_1 + \theta_2)/2}{dt} = k_\theta [\theta(t) - \theta_{\text{out}}(t)] \quad (3.84)$$

$$\frac{d(\theta_1 - \theta_2)/2}{dt} = k_r [r(t) - r_{\text{out}}(t)] \quad (3.85)$$

Recognize that $(\theta_1 + \theta_2)/2$ is the output signal phase— $\theta_{\text{out}}(t)$ and $r_{\text{max}} \cos(\theta_1 - \theta_2)/2$ is the output signal amplitude— $r_{\text{out}}(t)$, the control equations, surprisingly, correspond to the decoupled phase and amplitude, which may be solved separately. The phase of the loop output can be solved analytically; that is,

$$\theta_{\text{out}}(t) = k_\theta e^{-k_\theta t} \int_0^t \theta(t) e^{k_\theta t} dt \quad (3.86)$$

The amplitude of the loop output is given by

$$\frac{d\psi_{\text{out}}(t)}{dt} = k_r r_{\max} \left[\cos \psi(t) - \cos \psi_{\text{out}}(t) \right] \quad (3.87)$$

and may be solved by numerical methods. It can be shown that when the input is a sinusoidal signal with angular frequency ω_0 , the system output has the phase error of

$$\Delta\theta = \frac{\omega_0}{k_\theta} \quad (3.88)$$

which is the same as in a PLL; increasing the loop gain will decrease the tracking error, while raising the issue of stability. The tracking error may be reduced and the sensitivity may be enhanced without increasing the loop gain by using integrators.

The feedback loop in a VLL allows for the automatic compensation of the path imbalance between the two VCO branches. A change in the amplitude or phase response of either amplifier branch gives rise to a change in the phase of both VCO output signals, which eventually results in a reduced imbalance. One major disadvantage of VLL (similar to CALLUM as discussed before) is its sensitivity to the input signal magnitude. Also, the loop fails in lock when the input reference reaches zero amplitude.

References

- [1] Casadevall, F. J., and J. J. Olmos, "On the Behavior of the LINC Transmitter," in *Proc. 40th IEEE Veh. Technol. Conf.*, Orlando, FL, May 6–9, 1990, pp. 29–34.
- [2] Casadevall, F. J., and A. Valdovinos, "Performance Analysis of QAM Modulations Applied to the LINC Transmitter," *IEEE Trans. Veh. Technol.*, Vol. 42, No. 4, November 1993, pp. 399–406.
- [3] Tomisato, S., K. Chiba, and K. Murota, "Phase Error Free LINC Modulator," *Electr. Lett.*, Vol. 25, No. 9, April 27, 1989, pp. 576–577.
- [4] Olson, S. A., and R. E. Stengel, "LINC Imbalance Correction Using Baseband Preconditioning," *1999 IEEE Radio and Wireless Conf. (RAWCON 99)*, Denver, CO, August 1999, pp. 179–182.
- [5] Stengel, R. E., and S. A. Olson, "Method and Apparatus Utilizing a Compensated Multiple Output Signal Source," U.S. Patent, 5,901,346, May 4, 1999.
- [6] Faulker, M., T. Mattsson, and W. Yates, "Automatic Adjustment of Quadrature Modulators," *Electron. Lett.*, Vol. 27, No. 3, January 1991, pp. 214–216.

- [7] Zhang, X., and L. E. Larson, "Gain and Phase Error Free LINC Transmitter," *IEEE Trans. Veh. Technol.*, Vol. 49, No. 5, September 2000, pp. 1986–1994.
- [8] Couch, L.W., *Digital and Analog Communication Systems*, 5th ed., Upper Saddle River, NJ: Prentice Hall, 1997.
- [9] Zhang, X., et al., "Gain/Phase Imbalance Minimization Technique for LINC Transmitter," *IEEE Trans. MTT*, Vol. 49, No. 12, December 2001, pp. 2507–2516.
- [10] Sundstrom, L., "Automatic Adjustment of Gain and Phase Imbalances in LINC Transmitters," *Elect. Lett.*, Vol. 31, No. 3, February 2, 1995, pp. 155–156.
- [11] Fletcher, R., *Practical Methods of Optimization*, 2nd edition, New York: John Wiley & Sons, 1987.
- [12] Ampem-Darko, S., and H. S. Al-Raweshidy, "Gain/Phase Imbalance Cancellation Technique in LINC Transmitters," *Elect. Lett.*, Vol. 34, No. 22, October 1998, pp. 2093–2094.
- [13] Kolanek, J. C., J. J. Shynk and B. Baseghi, "Closed Loop Calibration for an Amplitude Reconstruction Amplifier," U.S. Patent, 6,215,354, April 10, 2001.
- [14] Chandrasekaran, R., et al., "Adaptive Algorithm for Calibrating a LINC Amplifier," *RAWCON'01*, August 2001.
- [15] Haykin, S., *Adaptive Filter Theory*, 3rd ed., Upper Saddle River, NJ: Prentice Hall, 1996.
- [16] Senderowicz, D., D. A. Hodges, and P. R. Gray, "An NMOS Integrated Vector-Locked Loop," *IEEE International Symposium on Circuits and Systems*, Rome, Italy, May 10–12, 1982, Vol. III, pp. 1164–1167.
- [17] DaSavila, M. K., "Vector Locked Loop," U.S. Patent 5,105,168, April 14, 1992.
- [18] Bateman, A., "The Combined Analogue Locked Loop Universal Modulator (CALLUM)," in *Proc. IEEE Conf. Veh. Technol.*, Denver, CO, May 10–13, 1992, pp. 759–763.
- [19] Bateman, A., M. Li, and K. Y. Chan, "Implementation of the LINC Transmitter Using the Combined Analogue Locked Loop Universal Modulator (CALLUM)," *Seventh IEE European Conf. Mobile Personal Commun.*, September 13–15, 1993, pp. 31–37.
- [20] Chan, K., and A. Bateman, "Linear Modulators Based on RF Synthesis: Realization and Analysis," *IEEE Trans. Circuits Syst. I*, Vol. 42, No. 6, 1995, pp. 321–333.
- [21] Chan, K. Y., and A. Bateman, "Analysis and Topologies for High Efficiency Amplifiers Based on RF Synthesis," in *Proc. 6th International Conference on Radio Receivers and Associated Systems*, Bath, UK, September 26–27, 1995, pp. 212–217.
- [22] Jennings, D. J., A. Bateman, and J. P. McGeehan, "Adjacent Channel Power and Error-Vector Magnitude Performance of Reduced Complexity CALLUM Systems," *IEE Proc.-Commun.*, Vol. 146, No. 5, October 1999, pp. 297–302.
- [23] Jennings, D. J., and J. P. McGeehan, "A High-Efficiency RF Transmitter Using VCO-Derived Synthesis: CALLUM," *IEEE Trans. MTT*, Vol. 47, No. 6, June 1999, pp. 715–721.

4

Power-Combining and Efficiency-Enhancement Techniques

4.1 Introduction

The outphasing amplifier architecture allows for the use of amplifier components driven by constant envelope signals, which can lead to dramatically higher efficiency than linear amplifier operation. For example, Class E (switching-mode amplifiers) have been demonstrated with efficiency above 80% at 1 GHz; overdriven Class B, C, and F amplifiers can attain similar levels of efficiency. This has led to the expectation that outphasing amplifiers with excellent linearity and very high efficiency should be possible. This expectation, however, has not been fully met to date. In fact, there are fundamental limitations to the ultimate efficiency of outphasing amplifiers, although the desired result can be approximated in some circumstances. The limitation in real systems is related to the efficiency of power combining at the amplifier outputs.

In real circuits, a lossless power combiner used at the amplifier output produces significant interaction between its inputs, which leads to distortion and a reduction in efficiency. Alternatively, it is possible to employ power combiners that provide isolation between the input ports; in such systems the linearity of the amplifier can be preserved, but the combiners are not lossless, and there is significant efficiency reduction as a result. This chapter discusses design trade-offs for the output amplifiers and power combiner. The characteristics of power combiners are first reviewed. The choice of power amplifier to be selected for use in the outphasing architecture is then

discussed, with a particular focus on the distinction between switching and nonswitching amplifiers. Strategies for using lossless combiners with appropriate amplifier configurations in a way to maximize efficiency (such as Chireix combining and derivatives of it) are subsequently presented. The results show that while ideal efficiency is not obtainable over the full range of output power, it can be optimized at a particular power level. The best selection of power level to optimize the combiner depends on the characteristics of the signals to be amplified, particularly the probability distribution of output power versus time; accordingly, characteristics of representative signals used in wireless communications are described. The use of lossy combiners, which offer isolation between the component amplifiers, is also presented. It is shown that the inefficiency thus generated can be partially overcome with a simple power-recycling technique, which recovers a significant fraction of the power usually lost in the combiner.

4.2 Power-Combining Techniques for Outphasing Amplifiers

Modern microwave technology offers many opportunities for power combining, including transformers, hybrid couplers, baluns, transmission line combiners, Lange couplers, and Wilkinson combiners. A variety of these combiners are illustrated in Figure 4.1. Fundamental considerations applicable to all of these demonstrate that power combiners with isolated input ports are lossy, while combiners that are lossless have input ports that are not isolated from each other.

Since combiners are essentially linear circuits, their behavior at a given frequency can be described by an S matrix (or a corresponding Z or Y matrix). For a reciprocal three-port (where no active circuits, plasmas or external magnetic fields are employed), the S -matrix is of the form

$$S = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \quad (4.1)$$

Let

$$b_i = S_{ij}a_j \quad (4.2)$$

where a_j is the incident wave at port j and b_i is the reflected wave at port i ($i, j = 1, 2, 3$). We consider port 3 to be the output port of the hybrid

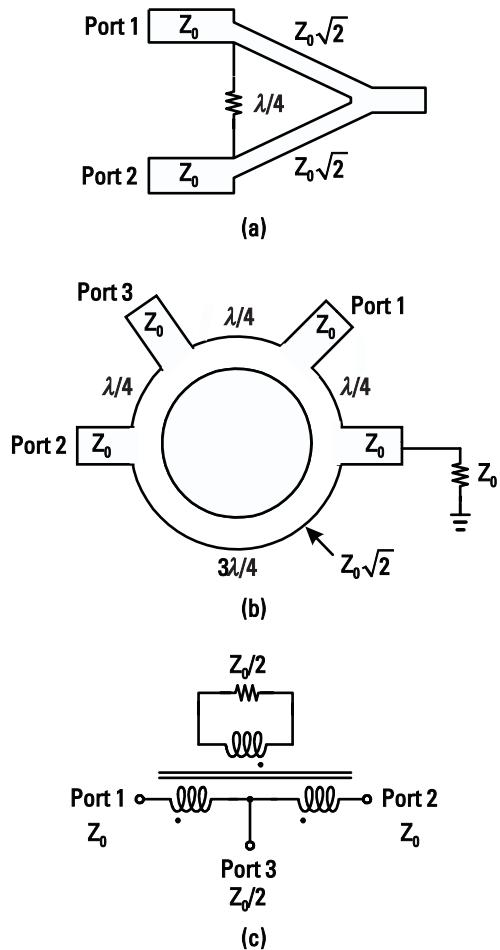


Figure 4.1 Schematic diagram of several different hybrid combiner structures:
 (a) Wilkinson combiner, (b) “Rat-Race” combiner, and (c) transformer hybrid.

and assume that it is matched to a chosen characteristic impedance Z_0 ; furthermore we consider the combiner/divider to have symmetric structure, so that there is equal power splitting between ports 1 and 2 for power incident on port 3. Then the S matrix can be further simplified to the form:

$$S_{\text{sym}} = \begin{bmatrix} \alpha & \beta & \gamma e^{j\phi_i} \\ \beta & \alpha e^{j\phi_j} & \gamma \\ \gamma e^{j\phi_i} & \gamma & 0 \end{bmatrix} \quad (4.3)$$

Here ϕ_i and ϕ_j are phase delays that can have different values according to the design (or choice of reference phase plane). The power combiner can be chosen to minimize interactions between the input ports. If we require that the input ports 1 and 2 are impedance-matched and isolated from one another, then α and β in (4.3) must equal zero. To provide for equal power splitting of a signal incident on port 3, it must also be true that $|\gamma| = 1/\sqrt{2}$. This leads to an S matrix given by:

$$S_{\text{isolated}} = \begin{bmatrix} 0 & 0 & e^{j\phi_1}/\sqrt{2} \\ 0 & 0 & e^{j\phi_2}/\sqrt{2} \\ e^{j\phi_1}/\sqrt{2} & e^{j\phi_2}/\sqrt{2} & 0 \end{bmatrix} \quad (4.4)$$

where ϕ_1 and ϕ_2 are design-dependent phase delays. This is the scattering matrix that describes Wilkinson combiners and combiners generated from hybrid couplers (lossless four ports) whose difference port is terminated in 50Ω . It is easy to see that power is not conserved in such a combiner under many circumstances. For example, if an input of unity amplitude is incident on port 1, then the output is given by a corresponding amplitude of $1/\sqrt{2}$ on port 3 (and no output on port 2), which corresponds to a lower output power than input power—an undesirable result!

An alternative design strategy is to use only lossless elements within the divider, so that the three-port network will also be lossless. Under such circumstances, the S matrix must be unitary; that is,

$$S \cdot S'^* = I \quad (4.5)$$

where I is the unit matrix. “ $'$ ” and “ $''$ ” denote “transpose” and “complex conjugate,” respectively. By systematically applying (4.5), the elements of the S matrix can be solved, with few undetermined parameters. In this case, the S matrix can then be simplified to the form:

$$S_{\text{lossless}} = \begin{bmatrix} e^{j\phi_0}/2 & -e^{j(\phi_0-\phi_1+\phi_2)}/2 & e^{j\phi_1}/\sqrt{2} \\ -e^{j(\phi_0-\phi_1+\phi_2)}/2 & -e^{j(\phi_0-2\phi_1+2\phi_2)}/2 & e^{j\phi_2}/\sqrt{2} \\ e^{j\phi_1}/\sqrt{2} & e^{j\phi_2}/\sqrt{2} & 0 \end{bmatrix} \quad (4.6)$$

Here again ϕ_0 , ϕ_1 , and ϕ_2 can have different values according to the design. The matrix shows that the input ports cannot be simultaneously matched to the characteristic impedance Z_0 ($S_{11} \neq 0$, $S_{22} \neq 0$); at the same time, there is considerable transmission of signals incident on port 1 to port 2 and vice versa ($S_{21} \neq 0$, $S_{12} \neq 0$). This S matrix describes many transformer-based, transmission line-based combiners and reactive combiners.

The behavior of the combiners in outphasing amplifiers is often clarified by reexpressing the S matrix in terms of even and odd modes incident on ports 1 and 2 (equivalent to common-mode and differential-mode excitations). Consider the usual case $\phi_1 = \phi_2$ (equal phase shifts from port 3 to ports 1 and 2). Using the basis functions

$$a_e = \frac{1}{\sqrt{2}}(a_1 + a_2) \quad (4.7)$$

$$a_o = \frac{1}{\sqrt{2}}(a_1 - a_2) \quad (4.8)$$

(and corresponding expressions for b_e , b_o), the S matrix representation can be converted to the following forms:

- For matched, isolated combiners:

$$\begin{bmatrix} b_e \\ b_o \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & e^{j\phi} \\ 0 & 0 & 0 \\ e^{j\phi} & 0 & 0 \end{bmatrix} \begin{bmatrix} a_e \\ a_o \\ a_3 \end{bmatrix} \quad (4.9)$$

- For lossless combiners:

$$\begin{bmatrix} b_e \\ b_o \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & e^{j\phi_2} \\ 0 & e^{j\phi_1} & 0 \\ e^{j\phi_2} & 0 & 0 \end{bmatrix} \begin{bmatrix} a_e \\ a_o \\ a_3 \end{bmatrix} \quad (4.10)$$

In the new basis, the corresponding expressions for the combiner operation appear in diagonal form. For lossless combiners, (4.10) expresses the relationship that, while even-mode excitations of the input ports are transmitted directly to the output without reflection, the odd mode excitations are completely reflected back to the input ports. For matched, isolated combiners, the reflection of the odd mode is suppressed. This suppression is easily understood as the effect of the input resistor in the Wilkinson divider, or the effect of the termination on the difference port of the hybrid coupler derivatives.

If we consider different phase choices for ϕ_1 and ϕ_2 , then more general matrices are found. By proper choice of phase delay (or phase reference) for ports 1 and 2, for example, the role of the even and odd modes in (4.10) can be interchanged.

Under conditions where lossless combining is used, the presence of the interaction term S_{12} complicates amplifier design. To relate to the most familiar basis for design, it is useful to convert S -parameter descriptions to Z -matrix descriptions of the two-port combiners, using

$$Z = (I - S)^{-1}(I + S) \quad (4.11)$$

For simplicity we may assume that the output port 3 is terminated in Z_0 (so $a_3 = 0$). Then the input characteristics of the combiner are described by a 2×2 impedance matrix:

For matched, isolated combiners

$$\begin{bmatrix} V_e \\ V_o \end{bmatrix} = Z_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I_e \\ I_o \end{bmatrix} \quad (4.12)$$

or equivalently

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = Z_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \quad (4.13)$$

For lossless combiners

$$\begin{bmatrix} V_e \\ V_o \end{bmatrix} = Z_0 \begin{bmatrix} 1 & 0 \\ 0 & -\cot \frac{\phi}{2} \end{bmatrix} \begin{bmatrix} I_e \\ I_o \end{bmatrix} \quad (4.14)$$

or equivalently

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = Z_0 \begin{bmatrix} 1 - \cot \frac{\phi}{2} & 1 + \cot \frac{\phi}{2} \\ 1 + \cot \frac{\phi}{2} & 1 - \cot \frac{\phi}{2} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \quad (4.15)$$

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \frac{1}{4Z_0} \begin{bmatrix} 1 - \tan \frac{\phi}{2} & 1 + \tan \frac{\phi}{2} \\ 1 + \tan \frac{\phi}{2} & 1 - \tan \frac{\phi}{2} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad (4.16)$$

Here V_1 , V_2 and I_1 , I_2 are the port voltages and currents, and V_e , V_o , and I_e , I_o are the corresponding even- and odd-mode linear combinations. If the combiner design has $\phi = 0$, from (4.15) then the impedance matrix is not defined, and the admittance matrix should be used. In such circumstances, the odd-mode impedance tends to infinity, corresponding to complete reflection.

The cross-impedance term Z_{12} appearing in the matrix for the lossless combiners expresses the coupling between amplifier outputs dictated by the combiner, in the form

$$V_1 = Z_{11}I_1 + Z_{12}I_2 \quad (4.17)$$

$$V_2 = Z_{21}I_1 + Z_{22}I_2 \quad (4.18)$$

In a conventional transmission-line combiner, the phase angle $\phi = 90^\circ$, so that $Z_{12} = Z_{11}$. The effect of the coupling is a strong “load-pulling” for the amplifiers. If the output currents I_1 and I_2 of the two amplifiers in the outphasing system have the relationship $I_2 = I_1 e^{j\zeta}$, then the effective load impedance experienced by amplifier “1” is

$$Z_{\text{out}1} = \frac{V_1}{I_1} \quad (4.19)$$

$$= Z_{11}(1 + e^{j\zeta}) \quad (4.20)$$

Thus as the output power of the outphasing amplifier changes and correspondingly ζ changes, the impedance seen by each amplifier also changes, with, in general, a change in both real and imaginary parts. As we will see, this leads to difficulties in maintaining high efficiency and high linearity over the different power output levels.

4.3 Amplifier Choices for Outphasing Systems

A variety of amplifier configurations can be considered for use in an outphasing power amplifier system, including Class A, B, and C (non-switching amplifiers), and Class E, F, and D (switching-mode amplifiers). The effects of the choice of power combining can be considerably different for the different configurations; as a result, the amplifier choice and the power combiner choice must be made together. A key difference is that in the conventional Class A, B, and C amplifiers, the transistor can be approximated as a current source, while for other classes of operation, the transistor is a voltage source, or a hybrid between the two extremes. Figure 4.2 shows representative transistor I-V characteristics. In region 1, the transistor output current is largely independent of output voltage (apart from a small slope of the curves, corresponding to a small value of output conductance).

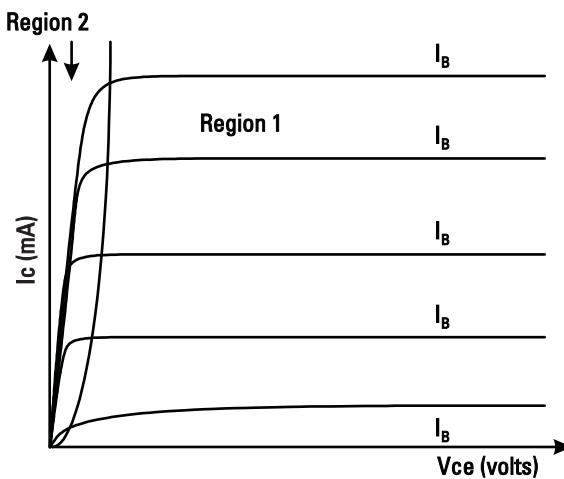


Figure 4.2 Bipolar transistor I-V curves, showing regions of constant current (region 1) and small output resistance (region 2).

When the input is below cutoff, the output current is also a current source, of value zero. On the other hand, in region 2, corresponding to transistor saturation [or the “linear region” of field effect transistors (FETs)] the transistor acts as a voltage source of value zero, with a small equivalent source resistance. Class A, B, and C amplifiers are customarily designed for use in region 1. In “overdriven” Class AB, B, and C amplifiers, during a portion of the RF cycle the transistor enters the saturation region, and thus acts as a voltage source during this interval. In Class E and F switching amplifiers, the transistor alternates between off and on states, in which it acts as a current source and a voltage source, respectively. In voltage-mode Class D amplifiers, two transistors are used in tandem with the result that their combination acts approximately like a voltage source at all times.

If the outphasing architecture is implemented with combiners having isolated input ports, then amplifier choice is relatively unrestricted. However, if lossless power combiners are chosen and thus the overall output impedance of each component amplifier is established by the outputs of both amplifiers, then the output interactions must be carefully designed. For amplifiers that act as voltage sources (in region 2), the overall impedance presented to the amplifier by the power combiner must not go to zero for any possible phase difference between the two amplifiers. By contrast, with current sources at the transistor outputs (in region 1), the overall admittance must not go to zero.

4.4 Outphasing Amplifier Design Using Class A, B, and C Amplifiers

To maximize the efficiency of each component amplifier, it is necessary that the power dissipation within the output transistor (given by the product of transistor current and transistor voltage) be minimized. This is typically accomplished by amplifiers operating in the *Class B* or *Class C* mode of operation. Figure 4.3 illustrates, for example, the transistor output voltage and output current for a representative Class B amplifier. The current waveform is approximately half-sinusoidal—a waveform that contains the fundamental frequency as well as even harmonics. The output voltage

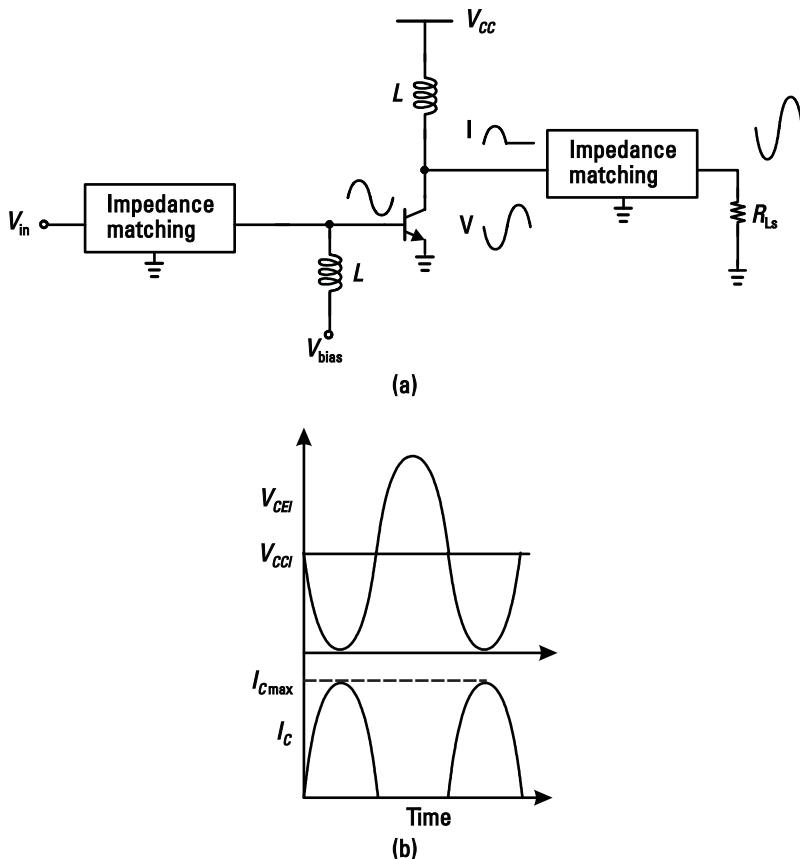


Figure 4.3 Class B power amplifier (a) schematic diagram, and (b) current and voltage waveforms.

contains only the fundamental frequency, since the load includes a short for all harmonics. The power dissipation within the transistor can be minimized by maximizing the voltage swing across the transistor (operating at maximum power) as well as by insuring that the phasing between output current and voltage is appropriate. By "appropriate," we mean that the current is zero when the voltage is maximum, and the current is highest when the output voltage is low. The (drain or collector) efficiency η as a function of the output voltage and current can be easily calculated.

If we set the collector current to

$$I(t) = \begin{cases} I_o \sin \omega_o t & \text{for } n\pi < \omega_o t < \pi + n\pi \\ 0 & \text{otherwise} \end{cases} \quad (4.21)$$

(This current waveform has an amplitude at the fundamental frequency ω_o of $I_o/2$) and assume

$$V_{CE}(t) = V_{dc} + V_o \sin(\omega_o t + \theta) \quad (4.22)$$

Then the mean values of current and voltage are given by:

$$\langle I(t) \rangle = \frac{I_o}{\pi} \quad (4.23)$$

$$\langle V(t) \rangle = V_{dc} \quad (4.24)$$

and the dc power dissipation is given by

$$\begin{aligned} P_{dc} &= \langle V \rangle \langle I \rangle \\ &= \frac{V_{dc} I_o}{\pi} \end{aligned} \quad (4.25)$$

and the RF power delivered to the load is given by

$$P_{RF} = \frac{1}{4} I_o V_o \cos \theta \quad (4.26)$$

and the efficiency of the amplifier is given by

$$\begin{aligned} \eta &= \frac{P_{RF}}{P_{dc}} \\ &= \frac{\pi}{4} \frac{V_o}{V_{dc}} \cos \theta \end{aligned} \quad (4.27)$$

The efficiency can be optimized by letting $\theta = 0$ through an appropriate choice of output load impedance, which should be real if the transistor acts as a current source with no shunt susceptance, and by operation at the highest possible value of V_o (generally Z_L is chosen so that $Z_L I_o / 2 = V_o = V_{dc} - V_{on}$, where V_{on} is the transistor “on” voltage). This leads to the well-known expression for the peak efficiency for the Class B amplifier of

$$\eta_{\max} = \frac{\pi}{4} \frac{V_{dc} - V_{on}}{V_{dc}} \quad (4.28)$$

The efficiency decreases, however, if the amplifier is operated at less than full power. Under these circumstances the efficiency is given by

$$\eta = \eta_{\max} \sqrt{\frac{P_{RF}}{P_{\max}}} \quad (4.29)$$

where P_{\max} is the maximum output power. In the outphasing amplifier using lossless combiners, the overall impedance seen by a component amplifier depends on the currents through both amplifiers. To determine the output voltage, we need to consider only the fundamental components of the currents, I_1 and I_2 (the overall current waveforms are half-sinusoidal as previously considered). Assume a conventional lossless transmission-line combiner, with $\phi = 90^\circ$, and Class B operation. The amplifier currents are assumed to be

$$I_1 = I_o e^{j\psi} \quad (4.30)$$

$$I_2 = I_o e^{-j\psi} \quad (4.31)$$

Using for the power combiner the relations from (4.15)

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = Z_L \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \quad (4.32)$$

The voltage amplitudes are given by

$$\begin{aligned} V_1 &= V_2 \\ &= 2Z_L I_o \cos \psi \end{aligned} \quad (4.33)$$

The RF output power of each amplifier is $Z_L I_o^2 \cos^2 \psi$, and its dc power consumption is $2V_{dc}/(\pi I_o)$. This leads to an efficiency of

$$\eta = \frac{\pi Z_L I_o \cos^2 \psi}{2 V_{dc}} \quad (4.34)$$

To avoid saturation of the amplifier, it is necessary to choose $2Z_L I_o < V_{dc} - V_{on}$. This leads to a best case efficiency for the Class B amplifier with lossless combining of

$$\eta_{max} = \frac{\pi}{4} \frac{V_{dc} - V_{on}}{V_{dc}} \cos^2 \psi \quad (4.35)$$

As the output power varies with the imposed modulation, $\cos^2 \psi$ varies, decreasing efficiency. For this amplifier,

$$\eta = \eta_{max} \frac{P_{RF}}{P_{max}} \quad (4.36)$$

This has the same relationship to P_{RF} as for a traditional Class A amplifier, except that η_{max} is higher, which is not very favorable for high efficiency over a broad range of output powers. The efficiency suffers because over much of the range, the voltage across the output amplifiers is less than the maximum possible value, and because the current and voltage waveforms are, in general, out-of-phase.

Superior efficiency can be obtained by “overdriving” the Class AB, B, or C amplifiers that make up the outphasing system. In this case, $2Z_L I_o$ is chosen to have a value greater than $V_{dc} - V_{on}$, so that during a portion of the RF cycle the amplifier is driven into saturation. With overdriven operation, the output voltage becomes constant, approximately independent of the input drive voltage or output current, with amplitude $V_{max} = V_{dc} - V_{on}$. The corresponding currents may have complex waveforms, containing multiple harmonics. The output load provides a short circuit at the harmonic frequencies, however, allowing the voltage to remain a pure fundamental tone. To analyze the situation, we assume

$$V_1 = V_{max} e^{j\psi} \quad (4.37)$$

$$V_2 = V_{max} e^{-j\psi} \quad (4.38)$$

Note that the phases of the output voltages are not necessarily exactly equal to the phases of the input voltage; this is one of the sources of distortion of this operating mode. With the transistors used as voltage sources, it is necessary to

choose a different configuration for the power combiner (avoiding the short circuit for odd modes provided by the transmission line combiner). With the choice $\phi = 0$, a combiner can be obtained with the following input relationship

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = Y_L \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad (4.39)$$

For the individual amplifiers, this leads to

$$I_1 = 2V_{\max} Y_L \cos \psi \quad (4.40)$$

The average current drawn from the power supply can be estimated to be

$$\langle I(t) \rangle = \frac{4}{\pi} V_{\max} Y_L \cos \psi \quad (4.41)$$

assuming the same relationship between the average and fundamental current component as for the underdriven Class B. This relationship is only approximate, and becomes progressively worse as the degree of overdrive increases. The RF output power and efficiency are then found to be

$$P_{\text{RF}} = V_{\max}^2 Y_L \cos^2 \psi \quad (4.42)$$

$$\eta = \frac{\pi}{4} \frac{V_{\max}}{V_{\text{dc}}} \cos \psi \quad (4.43)$$

Here $V_{\max} = V_{\text{dc}} - V_{\text{on}}$. This leads to a best case efficiency of

$$\eta_{\max} = \frac{\pi}{4} \frac{V_{\text{dc}} - V_{\text{on}}}{V_{\text{dc}}} \cos \psi \quad (4.44)$$

As the output power varies with the imposed modulation, $\cos \psi$ varies, decreasing efficiency. This variation, however, is significantly slower than the $\cos^2 \psi$ variation found in the previous case. For this amplifier,

$$\eta = \eta_{\max} \sqrt{\frac{P_{\text{RF}}}{P_{\max}}} \quad (4.45)$$

This reproduces the efficiency variation obtained with the individual (underdriven) Class B amplifier. In both cases, the efficiency benefits from the fact that the dc power supply current varies in accordance with the output power level. The principal cause of the efficiency reduction is the variable phase relationship between the voltage and current in the output devices.

It was first pointed out by Chireix [1] that the efficiency reduction at a given output power backoff could be countered by including an extra phase shift in the power-combining circuit. This is discussed in Section 4.5.

4.5 Chireix Power-Combining Technique

We saw in Section 4.4 that the time-varying relative phases between the two outputs of the power amplifiers caused a degradation in the overall efficiency. This is a serious practical problem with the classic outphasing power amplifier, since the whole purpose of the design is to improve the efficiency of the scheme.

We now consider an *overdriven* Class B amplifier, for which the output voltage waveform is approximately constrained by saturation of the transistor to have a constant amplitude. Harmonic shorts at the output insure that a sinusoidal waveform is maintained. The output voltages for the individual amplifiers are assumed to be

$$V_1 = V_{\max} e^{j\psi} \quad (4.46)$$

$$V_2 = V_{\max} e^{-j\psi} \quad (4.47)$$

A power combiner is chosen that has values of $\phi_0 = \phi_1 = \phi_2 = 0$. In addition, lossless circuit elements are added asymmetrically with respect to amplifiers 1 and 2, to provide a desired phase shift between the output current and voltages, as shown in Figure 4.4. Shunt susceptances B_a and $-B_a$

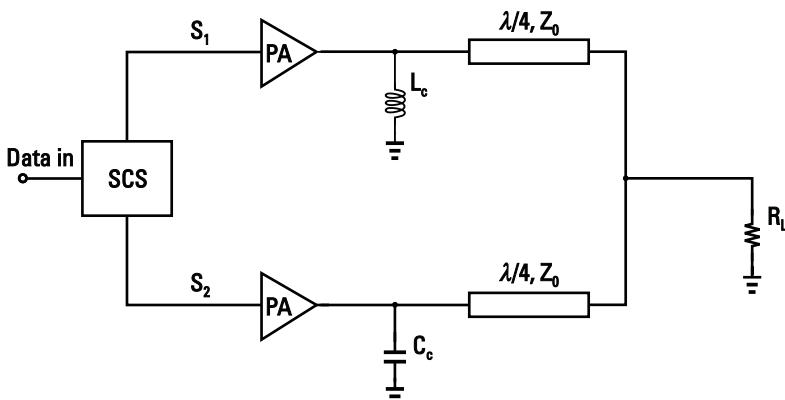


Figure 4.4 Chireix power-combining outphasing system with compensation shunt reactance [2].

(obtained with the capacitor C_c and inductor L_c , respectively) are added in parallel to the two sides. Now from (4.40)

$$I_1 = 2V_{\max} Y_L \cos \psi + jB_a V_{\max} e^{j\psi} \quad (4.48)$$

$$I_2 = 2V_{\max} Y_L \cos \psi - jB_a V_{\max} e^{-j\psi} \quad (4.49)$$

The RF power generated by amplifiers 1 and 2 is not changed by the additional susceptance, since the added current is out of phase with the corresponding amplifier voltage. Thus, as for the previous case,

$$\begin{aligned} P_{\text{RF1}} &= P_{\text{RF2}} \\ &= V_{\max}^2 Y_L \cos^2 \psi \end{aligned} \quad (4.50)$$

The current amplitude will change, however; that is,

$$|I_1|^2 = V_{\max}^2 (4Y_L^2 \cos^2 \psi + B_a^2 - 4Y_L B_a \cos \psi \sin \psi) \quad (4.51)$$

This can be minimized by choosing B_a such that $d|I_1|^2/dB_a = 0$, leading to

$$B_a = Y_L \sin 2\psi \quad (4.52)$$

This choice of B_a produces a phase shift to the current I_1 so that I_1 is in phase with V_1 .

In typical circuits, B_a has a fixed value, independent of output power level, so that the optimization can be done for only one particular choice of ψ , say $\psi = \psi_m$. Then

$$\begin{aligned} |I_1|^2 &= V_{\max}^2 (4Y_L^2 \cos^2 \psi + B_a^2 - 4Y_L B_a \cos \psi \sin \psi) \\ &= V_{\max}^2 Y_L^2 [(1 + \cos 2\psi)^2 + (\sin 2\psi_m - \sin 2\psi)^2] \end{aligned} \quad (4.53)$$

Correspondingly, the dc current consumption is

$$\langle I_1 \rangle = \frac{2}{\pi} V_{\max} Y_L \sqrt{(1 + \cos 2\psi)^2 + (\sin 2\psi_m - \sin 2\psi)^2} \quad (4.54)$$

which decreases due to the addition of the susceptances B_a and $-B_a$. The same results for output power and current consumption apply to amplifier 2. The overall result for efficiency is

$$\eta = \frac{\pi}{4} \frac{V_{\max}}{V_{dc}} \frac{1 + \cos 2\psi}{\sqrt{(1 + \cos 2\psi)^2 + (\sin 2\psi_m - \sin 2\psi)^2}} \quad (4.55)$$

When $\psi = \psi_m$ (corresponding to the optimal output power backoff),

$$\eta = \frac{\pi}{4} \frac{V_{\max}}{V_{dc}} \quad (4.56)$$

The amplifiers can therefore obtain the peak efficiency at a lower power than was obtained only at peak power, and the power where the peak efficiency occurs can be selected through the choice of ψ_m . At other power levels, the efficiency drops off in a manner described by (4.55). This is illustrated in Figure 4.5 for a variety of choices of ψ_m (or correspondingly, $B'_{\text{norm}} = B_a/Y_L$). The efficiency versus output power relationship expressed in these curves is, in general, much more favorable for system application than the original choice of $\psi_m = 0$. To understand how to optimize the choice of ψ_m , however, it is

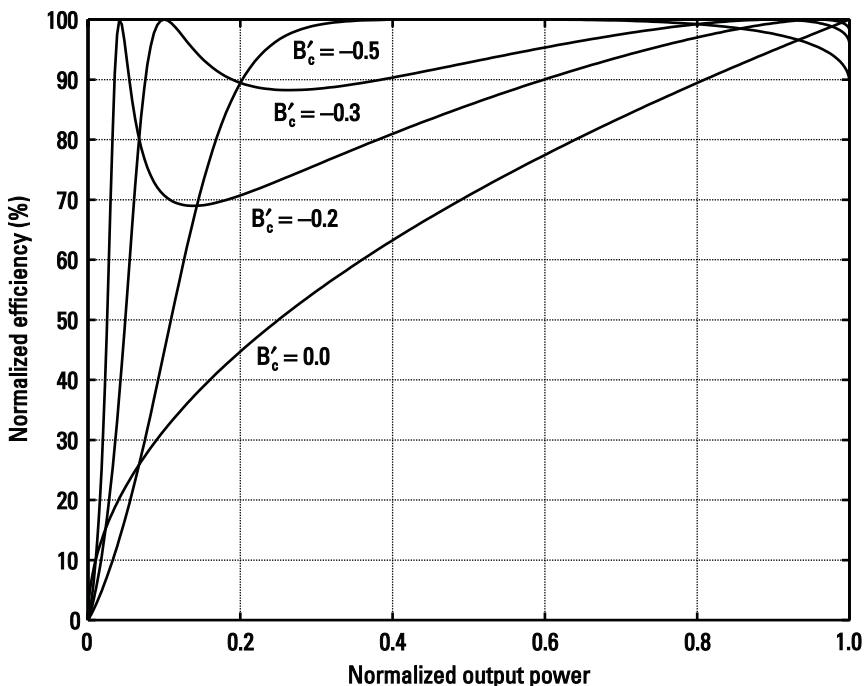


Figure 4.5 Normalized power efficiency versus normalized output power.

necessary to consider the power distribution for representative signals to be transmitted. This is undertaken in Section 4.8.

The Chireix technique is very attractive from the standpoint of efficiency. It has not led to date, however, to amplifiers with very high linearity. The causes for the reduction in linearity are related to: 1) the fact that Class AB, B, and C amplifiers in saturation are not exactly voltage sources—in reality, the output voltage is a function of the impedance seen at the amplifier output and the voltage varies in a nonlinear fashion—and 2) the phase of the output voltage does not bear a simple relationship to the phase of the input signal.

4.6 Combiner Design for Switching-Mode (Class D and Class E) Amplifiers

A potentially better approach for the implementation of the component amplifiers in the outphasing amplifier is to use switching-mode amplifiers. The efficiency of the amplifiers can in principle be excellent, since when the transistor voltage is high, the current is zero, and when the current is nonzero, the voltage is at its minimum level ($V_{\text{on}} \approx 0$). Unfortunately, there can be significant dissipation for the switching-mode amplifiers associated with the on-off transitions. A particularly, well-known dissipation mechanism is related to capacitance C_{out} at the transistor output node. If the output voltage is V_{sw} at the instant of switch closure, then an energy $C_{\text{out}} V_{\text{sw}}^2 / 2$ is dissipated in the on-resistance R_{on} of the transistor, during the ensuing transient discharge, independent of the value of R_{on} .

To minimize this dissipation mechanism, amplifier architectures have been developed for which “zero-voltage” switching is obtained; that is, $V_{\text{sw}} = 0$. For example, in the Class E amplifier shown in Figure 4.6, a tuned load resistance is chosen so that the voltage waveform has a characteristic shape leading to $V_{\text{sw}} = 0$ and $dV_{\text{sw}}/dt = 0$ at the instant of switch closure. With Class E amplifiers, efficiency above 80% has been demonstrated at 2 GHz, and above 60% at 10 GHz [2]. Unfortunately, the characteristics of the Class E amplifiers are not well-suited to outphasing systems with lossless power combining, since the zero-voltage switching feature is critically dependent on the phase of the load impedance. With the output power-dependent load-pulling described previously, this characteristic is lost. At the same time, the output voltage, current, and power are dependent on the load impedance, so that the constant envelope operation is also lost. Class E amplifiers are best suited for use with lossy combiners, discussed in Section 4.7.

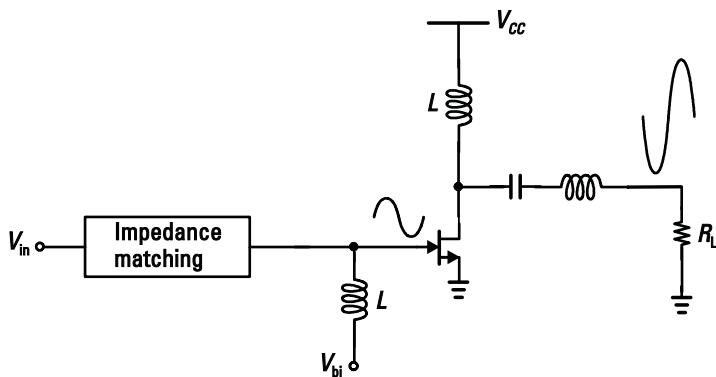


Figure 4.6 Class E amplifier schematic.

In the Class D amplifier (or more strictly, the voltage-mode Class D), the zero-voltage switching feature is absent, and as a result, its efficiency is not as high as for Class E (and grows progressively worse as the frequency is increased). Nonetheless this amplifier has advantages for outphasing implementation. As shown in Figure 4.7, the Class D amplifier employs two transistor switches that alternately connect the output to the positive and negative supply voltage [in this case, as in a complementary metal-oxide-silicon (CMOS) inverter, although other approaches are also possible].

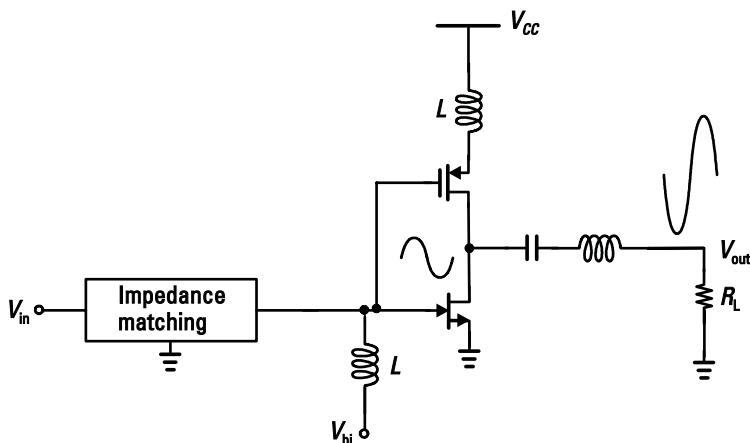


Figure 4.7 Class D amplifier schematic.

As a result, the amplifier acts as a voltage source, with a square-wave output. The load impedance is chosen so that there is nonzero conductance only at the fundamental frequency f_o (and the load provides an open circuit at all harmonic frequencies). Then the output current is a perfect sine wave at f_o .

In an outphasing amplifier, the phase of the current with respect to the voltage will vary according to the output power. In the Class D amplifier, however, the transistor power dissipation (apart from the on-off transitions) is near-zero, independent of the current phase angle. This differs from the overdriven Class B amplifier, which acts as a voltage source with a sinusoidal output waveform. In the Class D case, the current waveforms for the pull-up and pull-down switches are non-zero only when the switch voltage is near zero. When the phase of the load current waveform differs from that of the voltage, there is positive and negative current flow through the switches (current is drawn from the power supply, and returned back to it at a later time). For a phase difference of 90° (corresponding to zero net output power from the outphasing amplifier), the average current drawn from the supply is zero, as required for high-efficiency operation.

To assess the opportunities for outphasing amplifier implementation with Class D amplifiers, a more extensive analysis is required, taking into account realistic parasitic elements. This is provided in the following, for one implementation example, based on MOS transistor switches.

4.6.1 Analysis of MOSFET-Based Class D Outphasing Amplifier with Lossless Combining

The use of CMOS technology for Class D amplifiers is especially attractive, because of the availability of complementary switching devices. We now consider the following simplified equivalent circuit for CMOS-based Class D power amplifiers with the transmission line coupler shown in Figure 4.8. The Class D power amplifier is modeled as an ideal square-wave voltage generator with a series resistance R_s . The capacitance C_d models the drain parasitic capacitance of the MOS transistor. The LC tank rejects the harmonics other than the fundamental component ω_0 , and the loaded $Q_L = \omega_0 C_0 R_L$ is assumed.

The power amplifiers generate two outphased square waveforms with voltage swing from zero to V_{DD} . Taking the Fourier series expansion, and assuming 50% duty cycle, we have

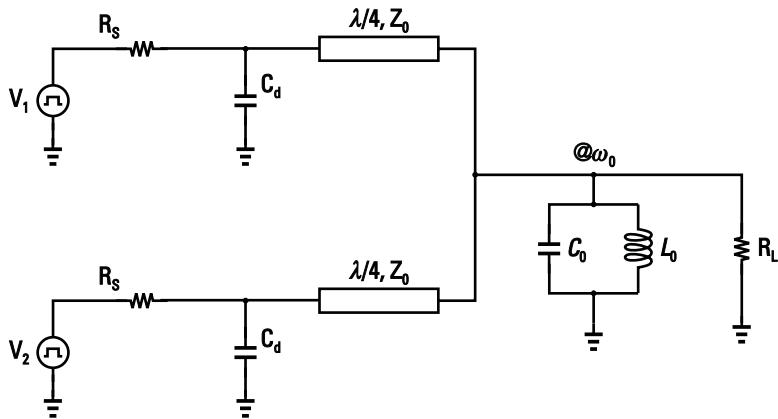


Figure 4.8 Simplified equivalent circuit of Class D outphasing power amplifier with the transmission line coupler. C_d is a parasitic capacitor at the output of the power devices.

$$V(\theta) = V_{DD} \left[\frac{1}{2} + \frac{2}{\pi} \sum_{n=1,3,5,\dots}^{\infty} \frac{1}{n} j^{n-1} \cos(n\theta) \right] \quad (4.57)$$

The above expression is translated into the two outphased signal component square waves according to Figure 1.17,

$$S_{1,2}(t) = \frac{V_{DD}}{2} + \frac{2}{\pi} V_{DD} \sum_{n=1,3,5,\dots}^{\infty} \frac{1}{n} j^{n-1} \cos[n(\omega_0 t \mp \psi)] \quad (4.58)$$

The dc constant of the voltage waveforms has no effect on the ac output and hence is neglected. $S_1(t)$ and $S_2(t)$ can be decomposed into the even and odd modes,

$$\begin{aligned} S_e(t) &= \frac{S_1(t) + S_2(t)}{2} \\ &= \frac{2}{\pi} V_{DD} \sum_{n=1,3,5,\dots}^{\infty} \frac{1}{n} j^{n-1} \cos(n\psi) \cos(n\omega_0 t) \end{aligned} \quad (4.59)$$

$$\begin{aligned} S_o(t) &= \frac{S_2(t) - S_1(t)}{2} \\ &= \frac{2}{\pi} V_{DD} \sum_{n=1,3,5,\dots}^{\infty} \frac{1}{n} j^{n-1} \sin(n\psi) \sin(n\omega_0 t) \end{aligned} \quad (4.60)$$

and with the following phasor notation on the n -order harmonic,

$$V_{e,n} = \frac{2}{\pi} V_{DD} \frac{1}{n} j^{n-1} \cos(n\psi) \quad (4.61)$$

$$V_{o,n} = \frac{2}{\pi} V_{DD} \frac{1}{n} j^{n-1} \sin(n\psi) \quad (4.62)$$

Note that the even- and odd-mode waveforms are not generally square-wave. The linear superposition theorem applies, so the circuit can also be rearranged according to even- and odd-mode operation and analyzed separately for the two cases. As the two odd-mode voltage generators drive the transmission line coupler 180° out of phase, due to the symmetry, the connection point between the two transmission lines shall act as a virtual ground. This situation is illustrated in Figure 4.9. The circuit can be further simplified because of the inverse impedance transform property of the quarter-wavelength transmission line. Looking into the transmission line from the capacitor side, the impedance is infinity. Thus, the circuit becomes a simple lowpass RC filter and the source current is given by

$$\begin{aligned} I_{o,n} &= V_{o,n} \frac{j n \omega_0 C_d}{1 + j n \omega_0 C_d} \\ &= \frac{2}{\pi} \frac{V_{DD}}{R_s} \frac{\delta \sin n\psi}{\sqrt{1 + n^2 \delta^2}} e^{j(n-1)\pi/2} e^{-j \arctan(n\delta)} \end{aligned} \quad (4.63)$$

where

$$\delta = \omega_0 C_d R_s \quad (4.64)$$

The time domain function is thus

$$\begin{aligned} I_{o,n}(t) &= \operatorname{Re}[I_o e^{j n \omega_0 t}] \\ &= \frac{2}{\pi} \frac{V_{DD}}{R_s} \frac{\delta \sin n\psi}{\sqrt{1 + n^2 \delta^2}} \cdot \sin[n\omega_0 t + \frac{n}{2}\pi - \arctan(n\delta)] \end{aligned} \quad (4.65)$$

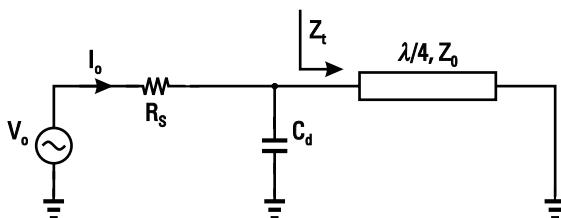


Figure 4.9 Odd-mode equivalent circuit of modified Chireix power-combining approach.

As expected, the odd-mode operation will not affect the output. However, power is consumed through the charge and discharge of the parasitic capacitor C_d .

For even-mode operation, the two voltage generators are operated in phase. Taking advantage of the symmetry, the circuit can be rearranged and simplified as in Figure 4.10. The load impedance is given by

$$\begin{aligned} Z_L|_{\omega=n\omega_0} &= \left(\frac{1}{2}j\omega C_0 + \frac{1}{2j\omega L_0} + \frac{1}{2R_L} \right)^{-1} \\ &= \frac{2R_L}{1+j\varepsilon} \end{aligned} \quad (4.66)$$

where

$$\varepsilon = Q_L \frac{n^2 - 1}{n} \quad (4.67)$$

After the quarter-wavelength transmission line, the load impedance becomes

$$\begin{aligned} Z_t &= \frac{Z_0^2}{Z_L} \\ &= R_t(1+j\varepsilon) \end{aligned} \quad (4.68)$$

where R_t is defined by

$$R_t = \frac{Z_0^2}{2R_L} \quad (4.69)$$

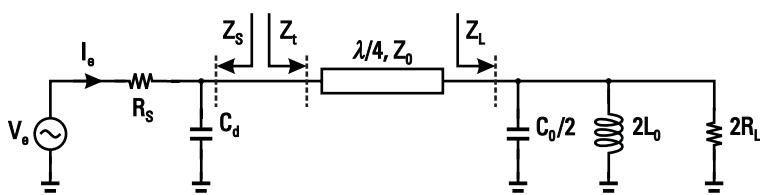


Figure 4.10 Even-mode equivalent circuit of modified Chireix power-combining approach.

The voltage generator together with the source resistance R_s and parasitic capacitance C_d can be regarded as a “black box.” By use of the Thevenin theorem, it is not difficult to obtain the voltage applied to the transmission line and then the final output

$$\begin{aligned} V_{\text{out}, n} &= V_{t,n} \frac{1 + \Gamma}{1 - \Gamma} j^{(n-2)} \\ &= -j^n \frac{Z_L}{Z_t} \frac{Z_t}{Z_t + Z_s} \frac{1}{1 + jn\delta} V_{e,n} \end{aligned} \quad (4.70)$$

The fundamental component is solved as

$$V_{\text{out}}(t) = \frac{2}{\pi} \frac{Z_0}{R_t} \cos \psi \frac{1}{\sqrt{(1 + \Delta)^2 + \delta^2}} \sin \left(\omega_0 t - \arctan \frac{\delta}{1 + \Delta} \right) \quad (4.71)$$

where

$$\Delta = R_s/R_t \quad (4.72)$$

and the output power is given by

$$\begin{aligned} P_{\text{out}} &= \frac{1}{2} \frac{V_{\text{out}}^2}{R_L} \\ &= \frac{4}{\pi^2} \frac{V_{DD}^2}{R_t} \frac{1}{(1 + \Delta)^2 + \delta^2} \cos^2 \psi \end{aligned} \quad (4.73)$$

With $\delta \ll 1$, which corresponds to the case where the RC constant of the transistor “on” resistance and drain parasitic capacitance is much less than the cycle of operation, we have

$$P_{\text{out}} = \left(\frac{2}{\pi} V_{DD} \right)^2 \frac{R_t}{(R_t + R_s)^2} \cos^2 \psi \quad (4.74)$$

The above result can be explained as follows: The first term $(2V_{DD}/\pi)^2$ is the fundamental component of the square wave, the second $R_t/(R_t + R_s)^2$ is the output portion of the series connection, and the last term $(\cos^2 \psi)$ corresponds the outphasing angle, which indicates the amplitude modulation. For the higher-order harmonics, we have the maximum amplitude of the harmonics applied to the load

$$\left| \frac{V_{\text{out}, n}}{V_{\text{out}, 1}} \right|_{\max} = \frac{1 + \Delta}{Q_L(n^2 - 1)} \frac{\cos n\psi}{\cos \psi}$$

$$\approx \begin{cases} \frac{1}{n^2 Q_L} & \text{when } \psi \rightarrow 0; \\ \frac{1}{n Q_L} & \text{when } \psi \rightarrow \frac{\pi}{2} \end{cases} \quad (4.75)$$

The source current can also be calculated,

$$I_{e,n}(t) = \frac{2}{\pi} \frac{V_{DD}}{R_s} \frac{\cos n\psi}{n} \sqrt{\frac{(\Delta - n\epsilon\delta)^2 + n^2\delta^2}{(1 + \Delta - n\epsilon\delta)^2 + (\epsilon + n\delta)^2}} \times$$

$$\sin \left[n\omega_0 t + \frac{n\pi}{2} + \tan^{-1} \frac{n\delta}{\Delta - n\epsilon\delta} - \tan^{-1} \frac{\epsilon + n\delta}{1 + \Delta - n\epsilon\delta} \right] \quad (4.76)$$

Now the total dc power consumption can be computed. Note that due to the square waveform of the voltage generators, we have

$$P_{dc} = \frac{1}{T} \int_{-T/2}^{T/2} [S_1(t)I_1(t) + S_2(t)I_2(t)] dt$$

$$= \frac{1}{T} \sum_{n=1,3,5,\dots}^{\infty} \int_{-T/4-\Delta t}^{T/4-\Delta t} [I_{e,n}(t) - I_{o,n}(t)] dt$$

$$+ \frac{1}{T} \sum_{n=1,3,5,\dots}^{\infty} \int_{-T/4+\Delta t}^{T/4+\Delta t} [I_{e,n}(t) + I_{o,n}(t)] dt$$

$$= P_e + P_o \quad (4.77)$$

where

$$\Delta t = \frac{\psi}{\omega_0} \quad (4.78)$$

Substituting (4.58), (4.65), and (4.76) into (4.77) and with the assumption of $Q_L \gg 1$, total dc power consumption except for the fundamental component in the even-mode operation is calculated as follows:

$$P_b = P_o + \sum_{n=3,5,\dots}^{\infty} P_{e,n}$$

$$\begin{aligned}
&= 2C_d V_{DD}^2 f_0 \cdot \left[\tanh\left(\frac{\pi}{2\delta}\right) - \frac{4}{\pi} \cos^2 \psi \frac{\delta}{1 + \delta^2} \right] \\
&\approx 2C_d V_{DD}^2 f_0 \quad \text{when } \delta \ll 1
\end{aligned} \tag{4.79}$$

which happens to be the power consumption of the parasitic capacitors with $0 - V_{DD}$ voltage swing and frequency f_0 . The system efficiency is hence given by

$$\begin{aligned}
\eta &= \frac{P_{\text{out}}}{P_{e,1} + P_b} \\
&\approx \frac{\Delta \cos^2 \psi}{\Delta(1 + \Delta) \cos^2 \psi + \frac{\pi}{4}(1 + \Delta)^2 \delta}
\end{aligned} \tag{4.80}$$

Define η_0 as the peak efficiency when the two amplifier branches are operated in phase, we have

$$\begin{aligned}
\eta_0 &= \frac{\Delta}{\Delta(1 + \Delta) + \frac{\pi}{4}\delta(1 + \Delta)^2\delta} \\
&= \frac{R_t}{R_t + R_s} \frac{R_s}{R_s + \frac{\pi}{4}(R_t + R_s)\delta}
\end{aligned} \tag{4.81}$$

and the efficiency is given by

$$\frac{1}{\eta} \approx \frac{1}{\eta_0} + \frac{\pi(1 + \Delta)^2}{4\Delta} \delta \tan^2 \psi \tag{4.82}$$

Figure 4.11 shows the calculated overall efficiency as a function of the output power. The quantity Δ is chosen as $1/6.7$. The dashed line shows the efficiency characteristic of an ideal Class B power amplifier for comparison. Note that the efficiency of this approach is much better than the case of the outphasing amplifier hybrid-combining approach.

4.6.2 Simulation and Discussion of MOS-Based Class D with Lossless Combining

The Class D approach with lossless combining is suitable for a CMOS or GaAs FET implementation, due to the near-ideal switching nature of these

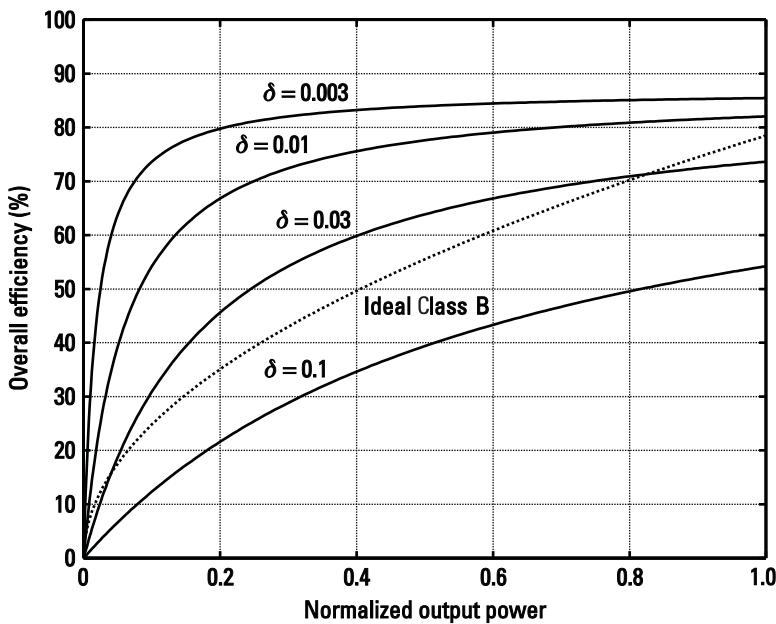


Figure 4.11 Calculated overall power efficiency versus normalized output power for various δ .

devices. The switching Class D power amplifiers used in the circuit would be much like a digital inverter, as illustrated in Figure 4.12. The driver stage is not shown. Two preamplifiers take the two signal components from the I/Q modulators and drive the CMOS inverters into the triode region as switches. When the input driving signal is near ground, the PMOS transistor is “on” and the NMOS is “off,” so the drain of the CMOS inverter is charged to V_{DD} through the PMOS. When the input signal is high, the inverse situation occurs and the drain is discharged to ground through the NMOS transistor.

A major problem of this approach is the parasitic capacitor C_d at the transistor drains, which consumes a considerable portion of power. With a 3.3-V supply voltage, as much as a 3-pF parasitic capacitance (around 0.3 pF/mm), and the circuit operated at 1 GHz, the power wasted by the capacitor is calculated to be around 110 mW, which could be 20% amount of the output power. The silicon on sapphire (SOS) CMOS technology employs an insulating sapphire substrate, and the transistor substrate is floating. With this technology, the parasitic capacitance is essentially eliminated, which makes it an attractive candidate for this modified

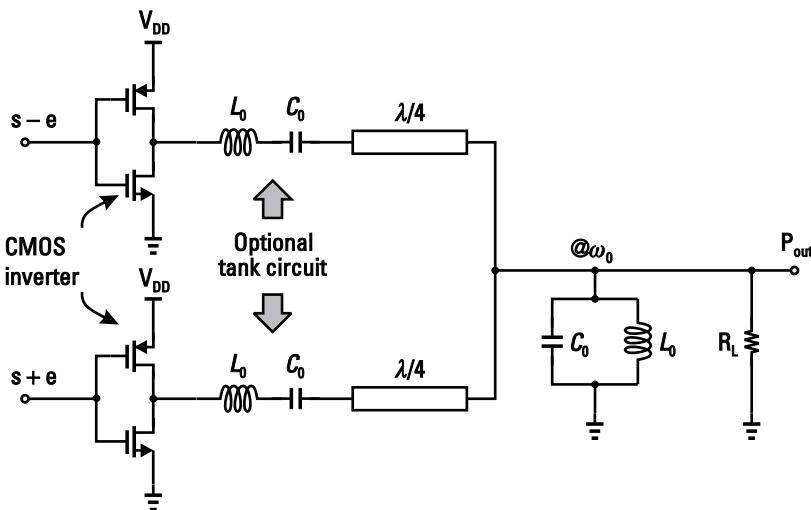


Figure 4.12 CMOS implementation of a Class D outphasing amplifier with transmission line coupler.

Chireix-combining method. Unfortunately, the SOS CMOS is not compatible with most bulk CMOS processes.

Figure 4.13 shows the simulated drain efficiency versus the normalized output power for a 0.6- μm CMOS technology where C_{ds} has been set to zero. The driving signal is nearly a square waveform and the driving power was not included. The p-type MOS (PMOS) transistor ($0.6 \mu\text{m} \times 4.5 \text{ mm}$ for each) was chosen to be three times larger than the n-type MOS (NMOS) ($0.6 \mu\text{m} \times 1.5 \text{ mm}$ for each) because of its relatively low mobility. The simulated maximum output power is 0.54W with drain efficiency of 70.6%. The operating frequency is 1 GHz. It can be seen from the plot that the efficiency stays above 50% until the output power is below 0.15W. For comparison, the drain-efficiency characteristics of an ideal Class B and Class A power amplifier are also displayed.

The SOS CMOS implementation of this new combining approach is attractive and could achieve high efficiency and high linearity. However, there are still some issues with this approach. First of all, the drain voltage swing is from zero to V_{DD} , instead of from zero to $2V_{DD}$ for Class B and C. So for the same size of transistor, the output power with this approach would be roughly reduced by a factor of four. Second, to reduce the transient loss while the transistor switches from “off” to “on” and vice versa, the transistor has to be driven very “hard” due to the large gate capacitor of MOS

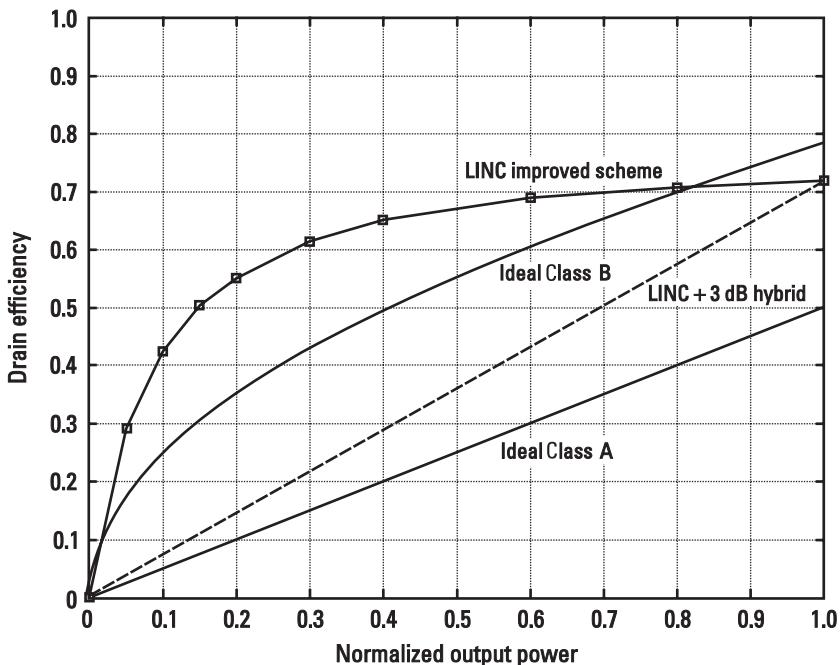


Figure 4.13 Simulated efficiency with $0.6\text{-}\mu\text{m}$ CMOS, with the parasitic drain capacitance turned off.

transistors. This will considerably increase the driving power, though it may be possible to decrease the driving requirement with added harmonics.

The first problem may be mitigated by increasing the supply voltage. With increased supply voltage, the drain efficiency may also be improved, since now the load can be transformed to a higher impedance to obtain the same amount of output power. However, this will cause a long-term reliability problem and breakdown, due to the relatively low breakdown voltage of the CMOS transistor. It may be possible to implement this approach with GaAs metal semiconductor FET (MESFET) technology, as shown in Figure 4.14. One problem, however, is the relatively high insertion loss of the transformer, typically 0.5–1 dB, in this frequency range.

4.7 Application of Lossy Power Combiners to Outphasing Power Amplifiers

The most straightforward approach to outphasing amplifier implementation is to use amplifiers designed for a *constant* load impedance, operated near

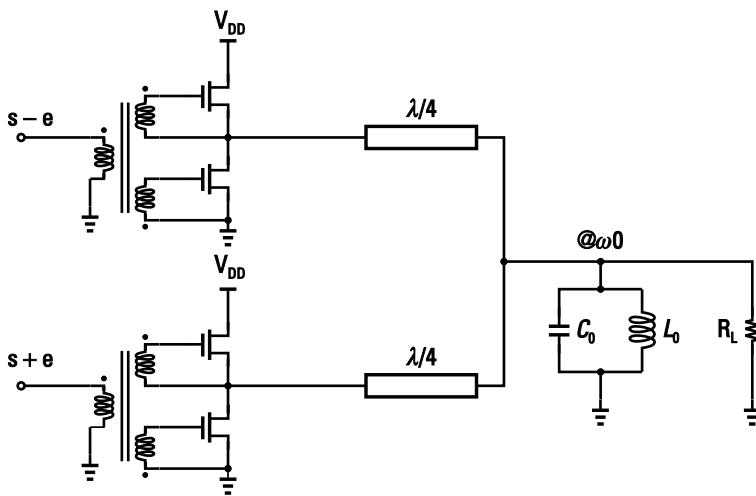


Figure 4.14 GaAs MESFET implementation of a Class D amplifier with transmission line coupler.

their saturated output power where their efficiency is maximum, together with a power combiner that provides isolation between the amplifier outputs (such as a Wilkinson combiner). This has the benefit that linearity of the outphasing system is enhanced because complications of signal-dependent loading of the amplifiers are avoided. As shown in Section 4.2, however, the power combining under these conditions is unavoidably lossy. Consider, for example, a combiner with an S matrix characterized by $\phi_1 = \phi_2$ [from (4.3)]. If the output wave amplitudes for amplifiers 1 and 2 are $a_o e^{j\psi}$ and $a_o e^{-j\psi}$, respectively, with output power $|a_o|^2/2$, the overall output power is easily seen to be

$$P = \frac{1}{2} \left| \frac{1}{\sqrt{2}} a_o (e^{j\psi} + e^{-j\psi}) \right|^2 \quad (4.83)$$

$$= |a_o|^2 \cos^2 \psi \quad (4.84)$$

If $\psi = 0$ (amplifiers in phase) the overall output power is equal to the sum of the individual amplifier powers, but if $\psi \neq 0$, the output power is reduced accordingly. If the efficiencies of the individual amplifiers is η_m , then the overall efficiency is

$$\eta = \eta_m \cos^2 \psi \quad (4.85)$$

$$= \eta_m \frac{P_{\text{RF}}}{P_{\text{max}}} \quad (4.86)$$

This is equivalent to the efficiency versus power relationship for the Class A amplifier, which exhibits a strong reduction in efficiency as the power decreases from its peak value.

An important example of this approach is the use of a conventional hybrid power combiner, as shown in Figure 4.15. In the ideal case, the hybrid completely isolates two amplifiers and each amplifier “sees” a $50\text{-}\Omega$ load. On the summing port of the hybrid, the in-phase signal components of the power amplifiers’ outputs add in power and the out-of-phase signal components cancel out. On the other hand, the opposite situation occurs on the difference port—the in-phase signal components cancel out and out-of-phase signal components add in power. Unfortunately, this portion of power is consumed at the $50\text{-}\Omega$ resistive load connected to the difference port and turns into waste heat, which degrades the overall power efficiency. This is the price

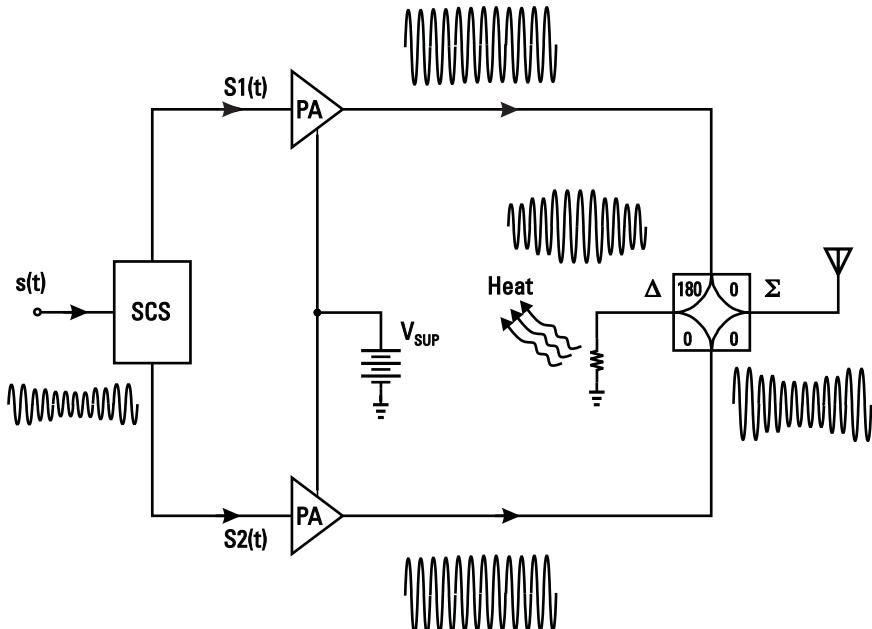


Figure 4.15 Outphasing power amplifier with hybrid power combining.

paid to prevent the load impedance variation. The amount of wasted power depends on the magnitude of the desired output signal.

There are two extreme cases. When the two amplifiers operate in-phase, twice the amount of each amplifier output power is delivered to the antenna, while no power is consumed on the load at the difference port. At the other extreme, when the two amplifiers operate completely 180° out of phase, all the power is wasted in the difference port, while no power is delivered to the antenna. Since the instantaneous combining efficiency is the ratio of the desired output power to the total power delivered by the two power amplifiers, the average combining efficiency is easily calculated by [3]

$$\eta_c = \int_0^1 \rho(r) r^2 dr \quad (4.87)$$

where r is the normalized transmitted signal magnitude, and $\rho(r)$ is the probability density function of the modulation. As a matter of fact, the wasted power is actually the quadrature signal— $e(t)$ —portion of the total transmitted power. The wasted power degrades the overall power efficiency, especially for high-level modulations, in which the signal experiences a wide range of variation of the power levels. For these modulations, the efficiency-power relationships from the lossless power-combining and Chireix-combining technique can provide advantages. To assess the relative merits of the different approaches, however, it is necessary to consider the distribution of power in various waveforms. This is considered in Section 4.8.

4.8 Probability Distribution of Output Power and Its Impact on Efficiency

For the different outphasing amplifier implementations, we have shown that the efficiency η varies with the instantaneous output power level. It is possible to calculate the average efficiency obtained with a given modulated waveform, given the probability density of output power for the modulation scheme. The average efficiency of the power amplifier can be determined from

$$\eta_{ave} = \frac{\langle P_{out} \rangle}{\langle P_{dc} \rangle} \quad (4.88)$$

where $\langle P_{out} \rangle$ and $\langle P_{dc} \rangle$ are the RF output power and dc power drawn from the supply averaged over time. These, in turn, can be calculated from

$$\langle P_{\text{out}} \rangle = \int P_{\text{out}} \rho(P_{\text{out}}) dP_{\text{out}} \quad (4.89)$$

$$\begin{aligned} \langle P_{\text{dc}} \rangle &= \int P_{\text{dc}} \rho(P_{\text{out}}) dP_{\text{out}} \\ &= \frac{P_{\text{out}}}{\eta(P_{\text{out}})} \rho(P_{\text{out}}) dP_{\text{out}} \end{aligned} \quad (4.90)$$

where $\rho(P_{\text{out}})$ is the probability of output power P_{out} . Thus the overall efficiency is

$$\eta = \frac{\int P_{\text{out}} \rho(P_{\text{out}}) dP_{\text{out}}}{\int \frac{P_{\text{out}} \rho(P_{\text{out}})}{\eta(P_{\text{out}})} dP_{\text{out}}} \quad (4.91)$$

The probability density function of the transmitted signal $\rho(P_{\text{out}})$ is required to calculate the combining efficiency, which is generally unavailable in analytical form. Hence the computation is based on the histogram obtained from simulation of the waveform over a long period of time. Figure 4.16 shows the simulated power probability density functions for various modulations. The modulations under investigation are QPSK, OQPSK, $\pi/4$ -DQPSK, 16-QAM, and 64-QAM. The baseband pulse shaping is the square-root raised cosine filtering, with a roll-off factor α varying from 0.1 to 1.0. Obviously the probability density functions are highly dependent on both the modulation and baseband filtering. For high-level modulations (i.e., 16- and 64-QAM) the probability density concentrates around the low power level more than it does for the lower-level modulations; hence we expect a much lower average-to-peak power ratio or a lower combining efficiency. Note that for these two modulations, the shape of the probability density function barely changes with the roll-off factor. In practical applications, the signal in-phase and quadrature components may be offset by half the symbol period for 16-QAM and 64-QAM, as in the case of OQPSK.

For Class A based outphasing amplifiers using lossy combining, and for outphasing amplifiers with lossless combiners using underdriven Class A, AB, and B amplifiers, we have seen that

$$\eta = \eta_{\max} \frac{P_{\text{out}}}{P_{\max}} \quad (4.92)$$

from which, it is easy to show that

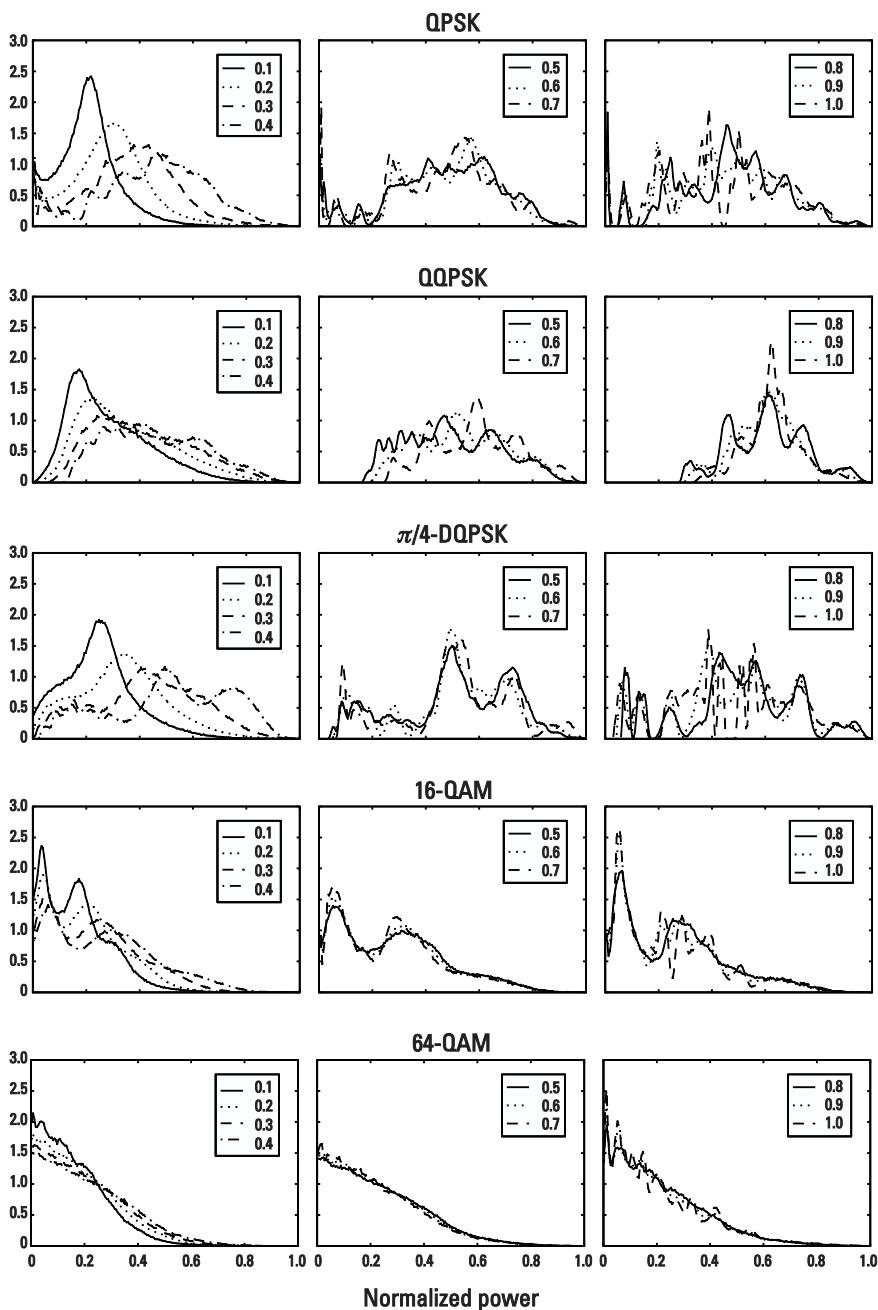


Figure 4.16 Power probability density function for various modulations. The legend indicates the roll-off factors of the square-root raised cosine filtering.

$$\eta_{\text{ave}} = \eta_{\max} \frac{\langle P_{\text{out}} \rangle}{P_{\max}} \quad (4.93)$$

This indicates that the average efficiency is degraded from the maximum efficiency (which is obtained at P_{\max}) by the PAP ratio of the waveform. In other words, the combining efficiency of the hybrid combiner is essentially equivalent to the average-to-peak power ratio of the signal.

Figure 4.17 shows the calculated combining efficiency as a function of the roll-off factor for various modulations, according to (4.87) and the simulated probability density function. As expected, for 16- and 64-QAM the combining efficiency is below 30% and varies little with the roll-off factor. For lower level modulations (e.g., QPSK, OQPSK, and $\pi/4$ -DQPSK) the combining efficiency is greater than 45% for a roll-off factor of 0.4 or larger and reaches as high as 62% for OQPSK. Note that the calculated combining efficiency does not account for the loss of hybrid combiner and efficiency of the power amplifier. To take these into account, the overall efficiency would be

$$\eta_{\text{overall}} = \eta_p \eta_c (1 - \rho_c) \quad (4.94)$$

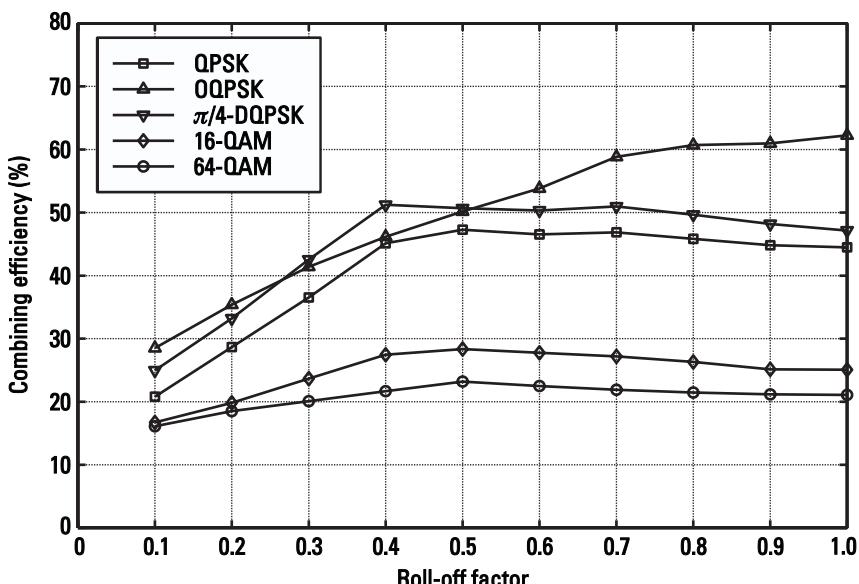


Figure 4.17 Combining efficiency as a function of roll-off factor for various modulations. Calculation is based on plot of Figure 4.16 and (4.87).

where η_p is the power amplifier drain efficiency, and ρ_c is the loss of the hybrid combiner. In fact, the combining efficiency would be a major degradation factor of the power efficiency for this combining approach. For simplicity later on, we will assume that the lossless hybrid combiner and the drain efficiency of the power amplifiers is 100%.

4.9 Power Recycling in Outphasing Amplifiers

The idea of a power recycling scheme is simple—replace the power-wasting resistive load in Figure 4.15 with a RF-dc converter to recover as much of the wasted power as possible back to the power supply and hence the overall power efficiency of the amplifier system is improved [4, 5]. The schematic diagram of a particular implementation of this method is illustrated in Figure 4.18. A 180° hybrid combiner is configured as the power splitter to divide the wasted power into two 180° outphased portions. These two signals are then fed to a high-speed Schottky diode pair through an impedance-matching network. The Schottky diodes rectify the RF waves and the dc components are withdrawn back to the power supply. A large value capacitor together with the RF chokes embedded in the power amplifiers shorts the harmonic current of the rectified wave to ground. Alternatively, an inductor might be used to reject the harmonics. The RF choke loops provide a dc return path for the circuit, if necessary. The matching network can be adjusted to optimize the power delivered. An optional isolator can be added between the hybrid combiner and power

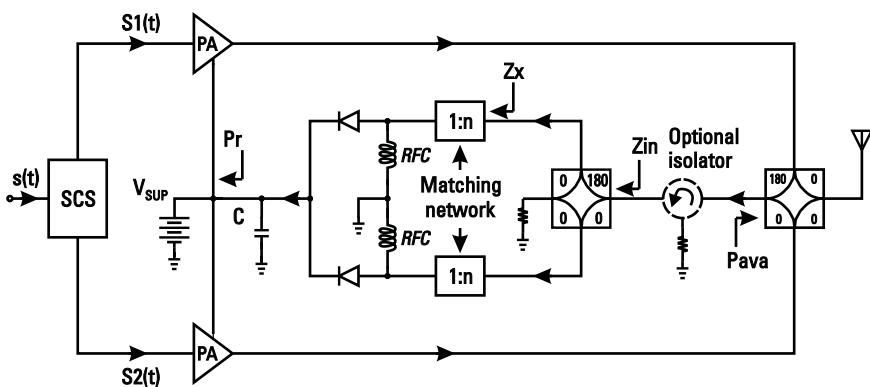


Figure 4.18 Outphasing power amplifier with power recycling.

splitter to improve the isolation. In a word, this power-recycling network is simply a microwave version of the full-wave rectifier.

Two figures of merit are important in the RF-dc conversion and need to be analyzed—the power efficiency of the recycling network and the isolation between two power amplifier branches in the presence of the recycling network. The isolation can be quantitatively analyzed by calculating the VSWR of the recycling network. The simplest case to analyze is the continuous wave, where the signal magnitude is a constant. In this case, the impedance-matching network can be adjusted to optimize the system performance. Intuitively, there might be an optimum point in which the recycling efficiency and the isolation and VSWR are maximized or minimized simultaneously. For linear modulations, the signal envelope varies and the recycling efficiency averages over the whole range of the signal power level, and the recycling efficiency and VSWR of the recycling network are traded off.

4.9.1 Analysis of the Power-Recycling Network for a Continuous-Wave Signal

To simplify the analysis, an ideal resistive model is assumed for the Schottky diode (i.e., the fixed “on-resistance” R_d in series with the built-in potential V_d , the infinite diode “off-resistance,” and a negligible shunt capacitance). All other components are assumed to be ideal. The analysis starts from the diode side. Since the 180° hybrid combiner is used as the power splitter, each diode conducts less than half the time and operates 180° out of phase. This situation is exactly the same as in the low-frequency full-wave rectifying circuit. Figure 4.19 illustrates the voltage and current waveforms for each Schottky diode. Note that each diode conducts at the angle of 2θ . The current through the upper diode is thus described by

$$I_d(t) = \begin{cases} \frac{V_{pk} \cos \omega_c t - (V_{sup} + V_d)}{R_d}; & \cos \omega_c t \geq \frac{(V_{sup} + V_d)}{V_{pk}} \\ 0; & \text{otherwise} \end{cases} \quad (4.95)$$

where V_{sup} is the power supply voltage, V_{pk} is the peak signal voltage applied to the diode, and ω_c is the carrier frequency of the signal. The diode conduction angle θ is determined by

$$\cos \theta = \frac{V_{sup} + V_d}{V_{pk}}; \quad 0 \leq \theta < \frac{\pi}{2} \quad (4.96)$$

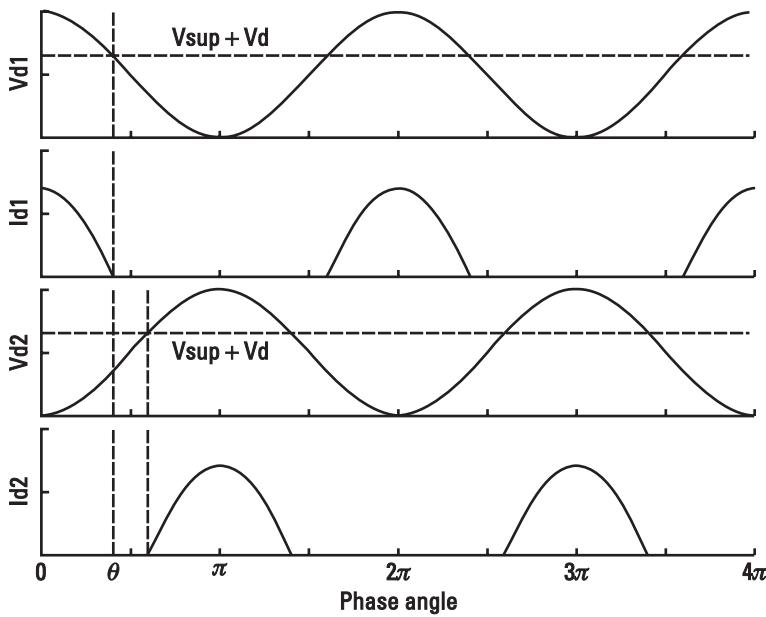


Figure 4.19 Diode voltage and current in the power-recycling network.

Note its symmetry with respect to the origin; the diode current can be expanded as the following Fourier series

$$I_d(t) = i_0 + \sum_{k=1}^{\infty} i_k \cos \omega_c t \quad (4.97)$$

where i_k is the k th-order harmonic of the diode current. The dc component per diode is thus given by

$$i_0 = \frac{V_{\text{sup}} + V_d}{\pi R_d} (\tan \theta - \theta) \quad (4.98)$$

The fundamental component and higher-order harmonics of the diode current are

$$i_{k,\text{upper}} = \frac{V_{\text{sup}} + V_d}{\pi R_d} \frac{1}{k \cos \theta} \left[\frac{\sin(k-1)\theta}{k-1} - \frac{\sin(k+1)\theta}{k+1} \right] \quad (4.99)$$

$$i_{k,\text{lower}} = \frac{V_{\text{sup}} + V_d}{\pi R_d} \frac{\cos k\pi}{k \cos \theta} \left[\frac{\sin(k-1)\theta}{k-1} - \frac{\sin(k+1)\theta}{k+1} \right] \quad (4.100)$$

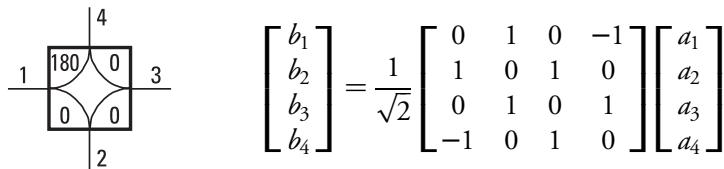
Obviously, the fundamental components and all odd-order harmonic current of the two diodes are 180° out of phase, and hence cancel out. Only the dc components and even-order harmonics are left. As we know that the RF “chokes” are usually embedded with the power amplifiers, a large value shunt capacitor may be sufficient to short the harmonic currents to the ground, as shown in Figure 4.18. A series inductor may be added to help reject the harmonic current out of the power supply. The recycled portion of the power is hence

$$P_r = 2i_0 V_{\text{sup}} \quad (4.101)$$

The RF-dc power conversion is a strongly nonlinear process. In such a case, the large signal impedance of the device is usually estimated by the fundamental component of the voltage and current waveforms. The fundamental component of the diode current at ω_c is thus

$$i_1 = \frac{V_{\text{sup}} + V_d}{\pi R_d} (\theta \sec \theta - \sin \theta) \quad (4.102)$$

The impedance-matching network scales the diode voltage and current seen from the hybrid splitter. We assume an impedance transformation ratio of $n : 1$ looking from the hybrid side to the diode side. The hybrid can also be considered as an impedance transformer. We use the scattering matrix to describe the hybrid and assume the following port convention [6]



where a_i and b_i , $i = 1, 2, 3, 4$ are the normalized incident and reflected voltage wave at each port, which are given by

$$a_i = \frac{V_{i,\text{inc}}}{\sqrt{Z_0}} \quad (4.103)$$

$$b_i = \frac{V_{i,\text{ref}}}{\sqrt{Z_0}} \quad (4.104)$$

where Z_0 is the characteristic impedance and assumed to be 50Ω . We now account for the fundamental components and calculate the recycling efficiency and the impedance looking into the recycling network. We have

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \quad (4.105)$$

here $a_3 = 0$, due to 50Ω resistive load at port 3. Assuming the impedance looking into the matching network from the hybrid splitter is Z_x , then

$$Z_x = n^2 \frac{V_{pk}}{i_1} \quad (4.106)$$

By applying (4.105), it is not difficult to obtain the following result

$$\begin{aligned} b_1 &= \frac{1}{\sqrt{2}}(a_2 - a_4) \\ &= \Gamma_x a_1 \end{aligned} \quad (4.107)$$

Thus, we have $\Gamma_{in} = \Gamma_x$, and $Z_{in} = Z_x$ —the input impedance of the recycling network is actually the impedance looking into the impedance matching network. So

$$\begin{aligned} Z_{in} &= n^2 \frac{V_{pk}}{i_1} \\ &= \frac{\pi n^2 R_d}{\theta - \sin \theta \cos \theta} \end{aligned} \quad (4.108)$$

and the reflection coefficient of the recycling network is

$$\Gamma_{in} = \frac{Z_{in} - Z_0}{Z_{in} + Z_0} \quad (4.109)$$

The VSWR is thus, according its standard definition

$$\text{VSWR} = \frac{1 + |\Gamma_{in}|}{1 - |\Gamma_{in}|} \quad (4.110)$$

The available power to the recycling network needs to be known to calculate the recycling efficiency. The input impedance of the recycling network varies with the diode conduction angle, which is determined by the power delivered to the recycling network. In practice the situation is complicated by the coupling between the variation of the load impedance and the power delivered by the amplifiers. For example, a load-impedance variation may drive a Class AB or Class B power amplifier into saturation or breakdown. A few power amplifiers that are sensitive to the load impedance and require careful tuning—like Class E amplifiers—may fail to work in this case. Furthermore, as stated previously, the load-impedance variation introduces an incomplete isolation between the two power amplifiers even with the hybrid combiner. Because of the load line match of the power amplifiers, the wideband quadrature signal reflected from the difference port of the hybrid is reflected to the summing port, though the hybrid provides an additional 3-dB isolation. The signal mixes with the desired output signal and creates the signal distortion. An isolator may be placed between the two hybrids to eliminate this effect. As shown in Section 4A.1, in this case looking from the hybrid combiner, the load is always matched to 50Ω , while looking from the hybrid splitter, the isolator acts like an ideal voltage source V_s in series with an internal resistance of 50Ω —the maximum power transfer theorem applies. Thus the source voltage is

$$\begin{aligned} V_s &= V_1 + Z_0 I_1 \\ &= \sqrt{2} n V_{\text{pk}} + \sqrt{2} \frac{Z_0 i_1}{n} \end{aligned} \quad (4.111)$$

where the scaling factor n results from the $1:n$ matching network and the $\sqrt{2}$ comes from the fact that the hybrid is a power-addition device. The available power to the recycling network is then

$$\begin{aligned} P_{\text{ava}} &= \frac{1}{8} \frac{V_s^2}{Z_0} \\ &= \frac{(V_{\text{sup}} + V_d)^2}{4Z_0} \left[n \sec \theta + \frac{Z_0}{n\pi R_d} (\theta \sec \theta - \sin \theta) \right]^2 \end{aligned} \quad (4.112)$$

and the recycling efficiency is

$$\eta_r = \frac{P_r}{P_{\text{ava}}} \quad (4.113)$$

The above expression implies that the recycling efficiency is a function of the diode conduction angle, which is determined by the available source power and the impedance-matching network.

Figures 4.20 and 4.21 show the 3D plots of the recycling efficiency and the VSWR as a function of the impedance transform ratio “ n ” of the matching network and the source available power “ P_{ava} .” The following typical diode parameters were chosen for the computation: $V_{sup} = 3V$, $V_d = 0.4V$, and $R_d = 10\Omega$. The contour plots are also shown beneath the mesh plots. It is clear that there is a close relationship between the optimum recycling efficiency and the lowest VSWR. There are two different cases—the constant power available to the recycling network and, more practically, the constant impedance transform ratio of the matching network. For the former case, the optimum recycling efficiency can be found by differentiating (4.113) with respect to the impedance transform ratio, under the constraint of the fixed available power; that is,

$$\frac{d\eta_r}{dn} = \frac{d\eta_r}{d\theta} \frac{d\theta}{dn} \quad (4.114)$$

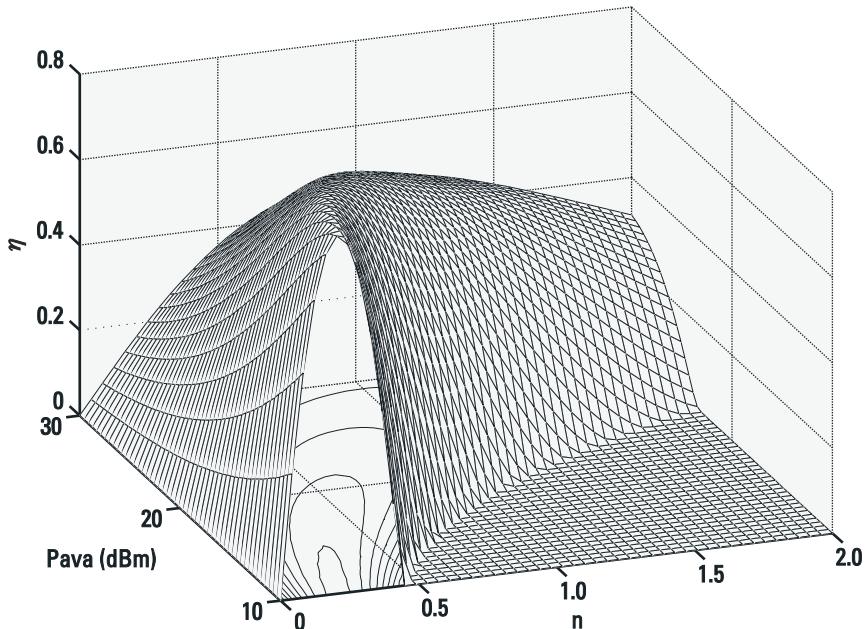


Figure 4.20 Recycling efficiency as a function of “ n ” and available power.

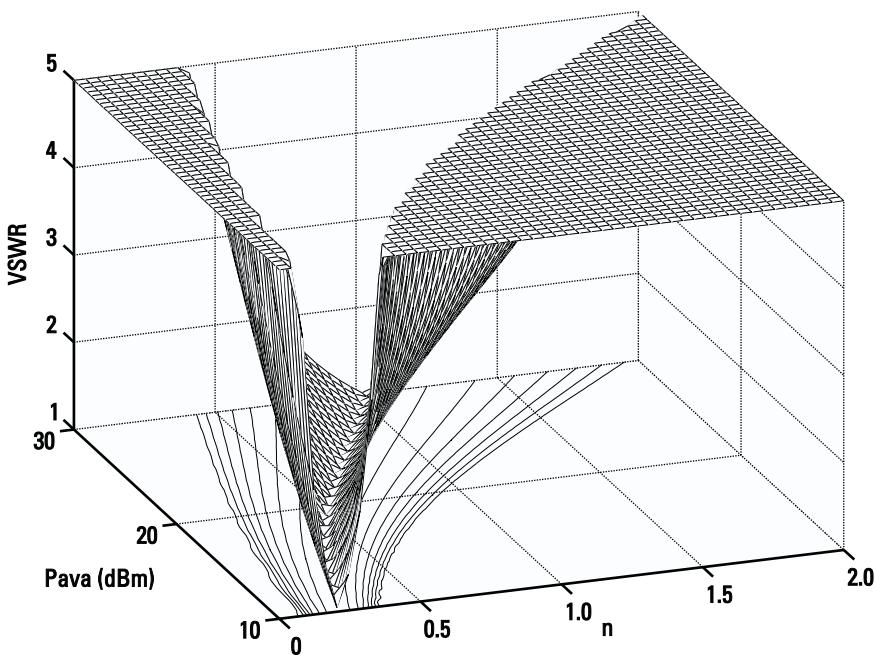


Figure 4.21 VSWR as a function of “*n*” and available power.

Equation (4.112) can be regarded as the implicit function of *n* and the diode conduction angle θ , since P_{ava} is fixed

$$\begin{aligned} F(n, \theta) &= \frac{1}{4Z_0} \left[nV_{\text{pk}}(\theta) + Z_0 \frac{i_1(\theta)}{n} \right]^2 - P_{\text{ava}} \\ &= 0 \end{aligned} \quad (4.115)$$

so the derivative of the diode conduction angle with respect to the impedance-transformation ratio $d\theta/dn$ can be found by

$$\begin{aligned} \frac{d\theta}{dn} &= - \frac{\frac{\partial F(n, \theta)}{\partial n}}{\frac{\partial F(n, \theta)}{\partial \theta}} \\ &= - \frac{V_{\text{pk}}(\theta) - \frac{Z_0}{n^2} i_1(\theta)}{n V'_{\text{pk}}(\theta) + \frac{Z_0}{n} i'_1(\theta)} \end{aligned} \quad (4.116)$$

Substitute (4.116) into (4.114) and finally we get

$$\frac{d\eta_r}{dn} = -\frac{2V_{\text{sup}}}{P_{\text{ava}}} \sin \theta \frac{V_{\text{pk}}(\theta) - \frac{Z_0}{n^2} i_1(\theta)}{n\pi R_d + \frac{Z_0}{n}(\theta + \sin \theta \cos \theta)} \quad (4.117)$$

The maximum recycling efficiency occurs when above expression is equal to zero, so we have

$$n^2 V_{\text{pk}}(\theta) - Z_0 i_1(\theta) = 0 \quad (4.118)$$

which is equivalent to saying that

$$\begin{aligned} Z_{\text{in}} &= n^2 \frac{V_{\text{pk}}}{i_1} \\ &= Z_0 \end{aligned} \quad (4.119)$$

This implies that the optimum VSWR of the recycling network is achieved along with the optimum recycling efficiency, when the available power to the recycling network is constant. This conclusion also applies to any arbitrary diode model and is the direct consequence of the maximum power transfer theorem, as proved in Section 4A.2. Now the maximum recycling efficiency can be calculated. From (4.112), (4.113), and (4.119), it can be shown that the maximum recycling efficiency as function of the diode conduction angle is given by

$$\begin{aligned} \eta_{r, \text{max}} &= \frac{V_{\text{sup}}}{V_{\text{sup}} + V_d} \frac{4(\tan \theta - \theta)}{2\theta - \sin 2\theta} \cos^2 \theta \\ &\approx \frac{4(\tan \theta - \theta)}{2\theta - \sin 2\theta} \cos^2 \theta \end{aligned} \quad (4.120)$$

assuming $V_{\text{sup}} \gg V_d$, which is a reasonable approximation in most practical applications. The diode conduction angle in the above expression, of course, corresponds to the case with optimum impedance matching. A figure of merit that indicates the capability of the recycling network can be defined as the recycling factor ϵ_r , which determines the diode conduction angle under the optimum impedance matching.

$$\begin{aligned}\epsilon_r &= \frac{2\pi R_d P_{ava}}{(V_{sup} + V_d)^2} \\ &= \frac{2\theta - \sin 2\theta}{\cos^2 \theta}\end{aligned}\quad (4.121)$$

Figure 4.22 shows the maximum recycling efficiency achievable and the recycling factor as a function of the diode conduction angle for the case of optimum impedance matching. Hence, the maximum achievable recycling efficiency is related to the recycling factor—the circuit parameters of the recycling network—through the diode conduction angle. For example, it can be seen that for recycling efficiency higher than 55%, the diode conduction angle must keep below 60°, or correspondingly the recycling factor must be chosen to be less than five. The diodes, power supply, and available power can be traded off according to (4.121). Also note an interesting result: The smaller the diode conduction angle, the more RF power is converted to dc and hence the higher the recycling efficiency. This reminds us of the fact that in classical power amplifier design, which can be considered as a dc-RF

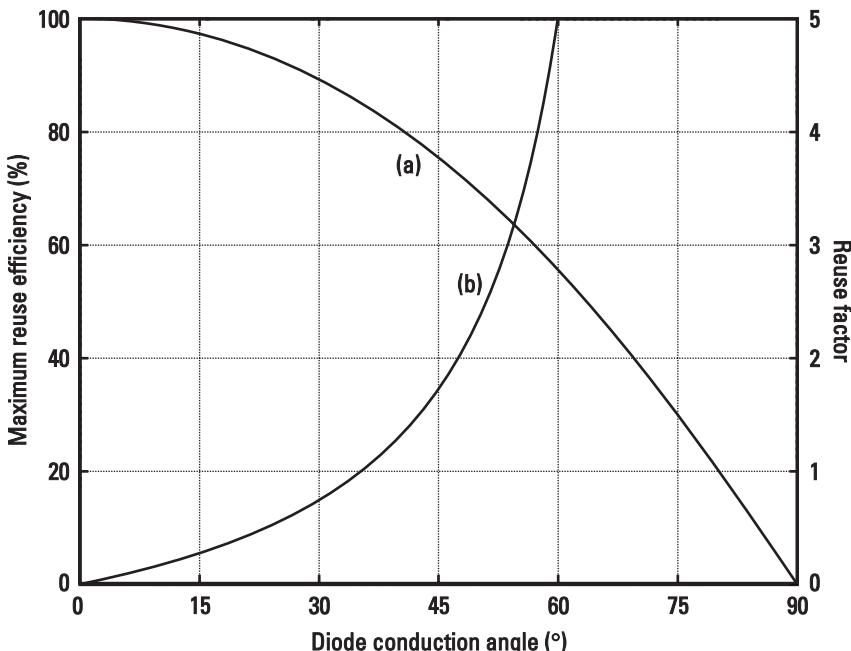


Figure 4.22 (a) Maximum recycling efficiency, and (b) recycling factor as a function of the diode conduction angle.

converter, the bias point of the transistor is lowered to reduce the transistor conduction angle and improve the drain efficiency. Figure 4.23 shows the maximum recycling efficiency as a function of the recycling factor.

Figures 4.24 and 4.25 show examples of the recycling efficiency and VSWR as a function of the impedance transform ratio for constant available power. The supply voltage varies from 3 and 4V to 5V. It is clear that the optimum recycling efficiency and VSWR coincide, which verifies the above analysis. This conclusion is also true for the realistic diode model, which is proved in Section 4A.2.

In practice, the matching network is usually fixed, and the available power to the recycling network varies with time—for example, in the case of linear modulated signals. The analysis shows that the maximum recycling efficiency generally does not occur along with the lowest VSWR. As an example, a set of calculated results are illustrated in Figures 4.26 and 4.27, which show the recycling efficiency and VSWR as a function of available power. The impedance transform ratio is fixed at $n = 0.3$, and the supply voltage varies between 3, 4, and 5V. It can be seen that the optimum recycling efficiency and VSWR are very close to each other, although they do not perfectly coincide.

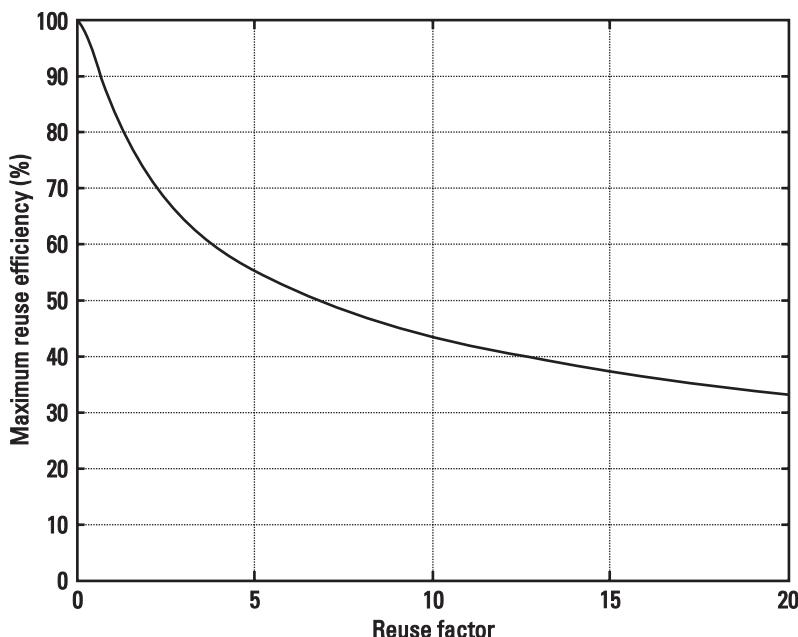


Figure 4.23 Maximum recycling efficiency as a function of the recycling factor.

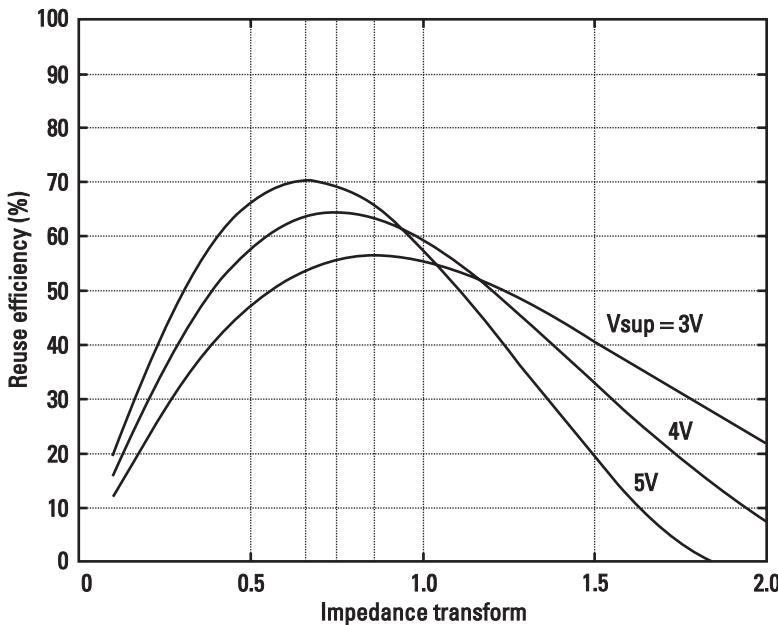


Figure 4.24 Recycling efficiency as a function of the impedance transform ratio for the constant available power.

Following similar procedures, the optimum efficiency can be determined by differentiating (4.113) with respect to the available power, under the constraint of fixed impedance transformation ratio:

$$\begin{aligned} \frac{d\eta_r}{dP_{ava}} &= \frac{2V_{sup}}{P_{ava}} \frac{di_0}{dP_{ava}} \\ &= \frac{2V_{sup}}{P_{ava}} \frac{\frac{di_0}{d\theta}}{\frac{dP_{ava}}{d\theta}} \end{aligned} \quad (4.122)$$

The analytical derivation is quite involved and we instead give the final results. It can be shown that the maximum recycling efficiency is given by

$$\begin{aligned} \eta_{r, \max} &= \frac{V_{sup}}{V_{sup} + V_d} \frac{3 \tan \theta + \theta(4\theta^2 \csc^2 \theta - 4\theta \tan \theta + \sec^2 \theta - 8)}{2\theta - 2 \tan \theta} \\ &\approx \frac{3 \tan \theta + \theta(4\theta^2 \csc^2 \theta - 4\theta \tan \theta + \sec^2 \theta - 8)}{2\theta - 2 \tan \theta} \end{aligned} \quad (4.123)$$

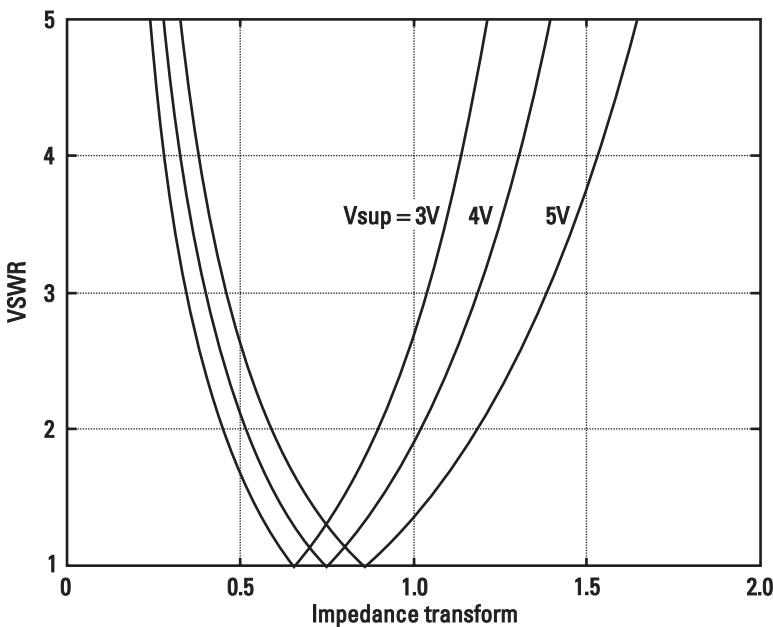


Figure 4.25 VSWR as a function of the impedance transform ratio for the constant available power.

and correspondingly, the input impedance of the recycling network with the optimum recycling efficiency is

$$Z_{in} = Z_0 \frac{8\theta^2 + 4\theta \cos 2\theta \tan \theta + 3 \cos 3\theta \sec \theta - 3}{2(\tan \theta - 2\theta)(2\theta - \sin 2\theta)} \quad (4.124)$$

The diode conduction angle in the above equations is determined by

$$\frac{Z_0}{n^2 \pi R_d} = \frac{4 \tan \theta - 8\theta}{8\theta^2 + 4\theta \cos 2\theta \tan \theta + 3 \cos 3\theta \sec \theta - 3} \quad (4.125)$$

Note that the diode conduction angle and hence the maximum recycling efficiency is independent of the supply voltage. Equations (4.123) and (4.124) imply that the conditions for the maximum recycling efficiency and for the lowest VSWR are generally different. In fact, they are fairly close to each other. This situation is illustrated in Figure 4.28, in which the maximum recycling efficiency and the correspondent VSWR are plotted as a function of the diode conduction angle. Obviously, in most cases the VSWR under the optimum efficiency is less than two, especially for the low-diode conduction angle.

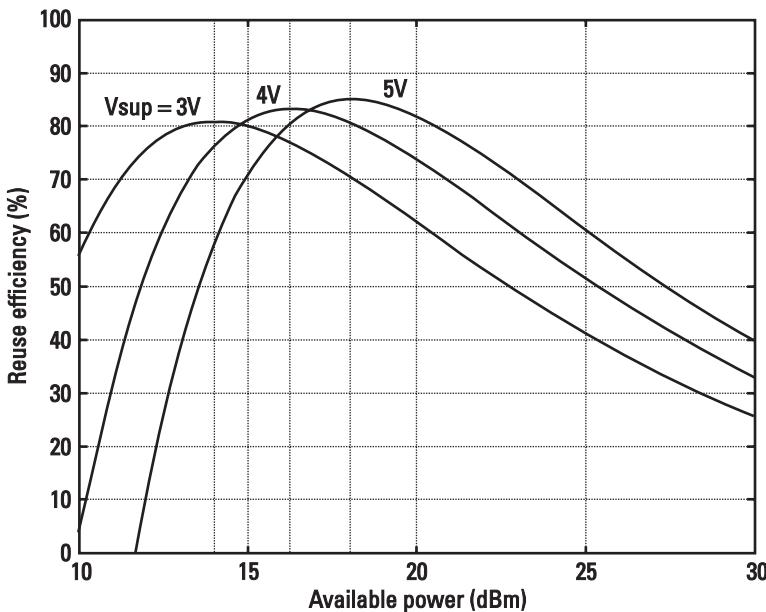


Figure 4.26 Recycling efficiency as a function of the available power for the constant impedance transform ratio.

4.9.2 Analysis and Discussion for Linear-Modulated Signals

For linear modulations, the signal magnitude and hence the available power to the recycling network exhibits a large variation. With the hybrid power-combining network, the overall power efficiency is the average-to-peak power ratio. By using the power-recycling network, a considerable portion of the wasted power is recovered back to the power supply, and the overall power efficiency is improved. In practice, the impedance-matching network is usually fixed. The total power delivered by the power amplifiers is a constant, and the power available to the recycling network is the quadrature signal portion of the total power, which varies with time. Assuming the recycling efficiency η_r , the instantaneous overall efficiency η_o is thus

$$\eta_o = \frac{p}{1 - (1 - p)\eta_r} \quad (4.126)$$

where p is the normalized output power to the antenna. Obviously without the recycling network, the overall efficiency becomes p ; on the other hand, if the recycling efficiency is 100%, the overall efficiency is 100% also.

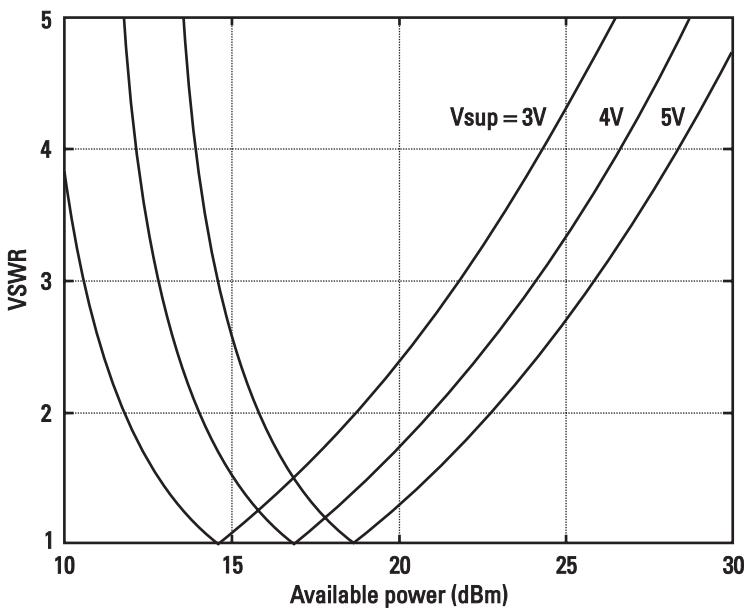


Figure 4.27 VSWR as a function of the available power for the constant impedance transform ratio.

Figure 4.29 displays the plot of the overall efficiency as a function of the output power for four different impedance transformation ratios. The following parameters are used for the calculation: $V_{\text{sup}} = 5\text{V}$, $V_d = 0.4\text{V}$ and $R_d = 5\Omega$. The dashed line corresponds to the case without power recycling, in which the power efficiency is equal to the normalized output power. With power recycling, the overall efficiency is enhanced. The solid lines show that the power efficiency improvement is dependent on the impedance matching network. The efficiency improvement near the low output power level is more critical than that in the high output power level. The reason is that when the output power is high, little power goes to the recycling network, and the efficiency is high anyway. Note that each solid line overlaps with the dashed line above a certain power level, depending on the matching network. Their intersection points correspond to where the diodes turns off, due to the fact that, beyond these points, less power is delivered to the recycling network and the voltage applied to the diodes is unable to overcome the supply voltage and turn on the diodes. The VSWR of the recycling network beyond these points is infinite, as shown in Figure 4.30. Proper choice of the impedance-matching network would result in a reasonably low VSWR (<2:1) across a wide range of the output power level.

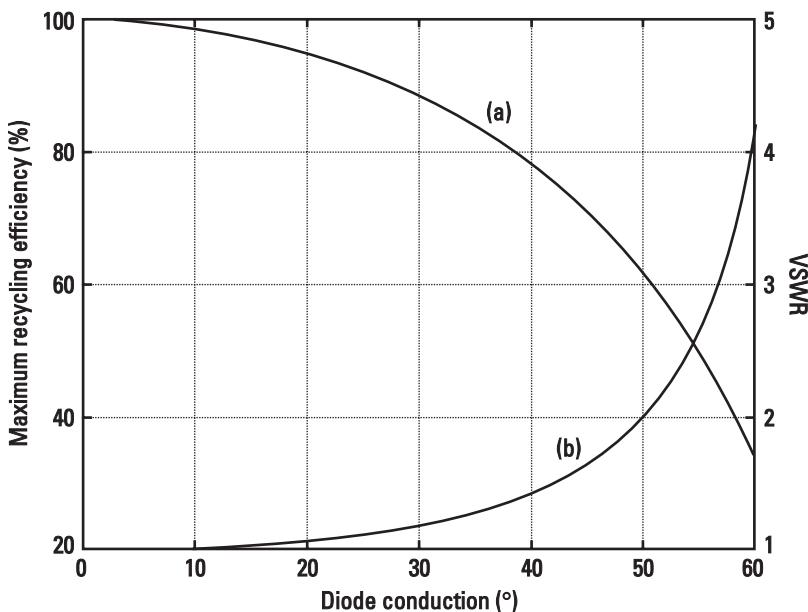


Figure 4.28 (a) Maximum recycling efficiency, and (b) corresponding VSWR as a function of the diode conduction angle.

For an even more strict requirement on VSWR, an isolator could be inserted between the recycling network and the hybrid combiner to completely eliminate the power reflected back to the power amplifiers. The disadvantage is that the insertion loss of the isolator degrades the effective power efficiency.

The average power efficiency for the linear modulated signal is calculated by integrating the instantaneous overall power efficiency weighted by the probability density over the output power; that is,

$$\bar{\eta}_o = \int_0^1 \frac{p}{1 - (1-p)\eta_r(p)} \rho(p) dp \quad (4.127)$$

where $\rho(p)$ is the power probability density function of the modulation calculated in Figure 4.16. The instantaneous overall power efficiency and hence the average power efficiency are strongly dependent on the impedance-matching network. At two extreme cases, when the transformation ratio “ n ” of the impedance-matching network tends to infinity, the voltage applied to the diodes is unable to turn them on most of the time, and the average power efficiency of the system is similar to the case without the power

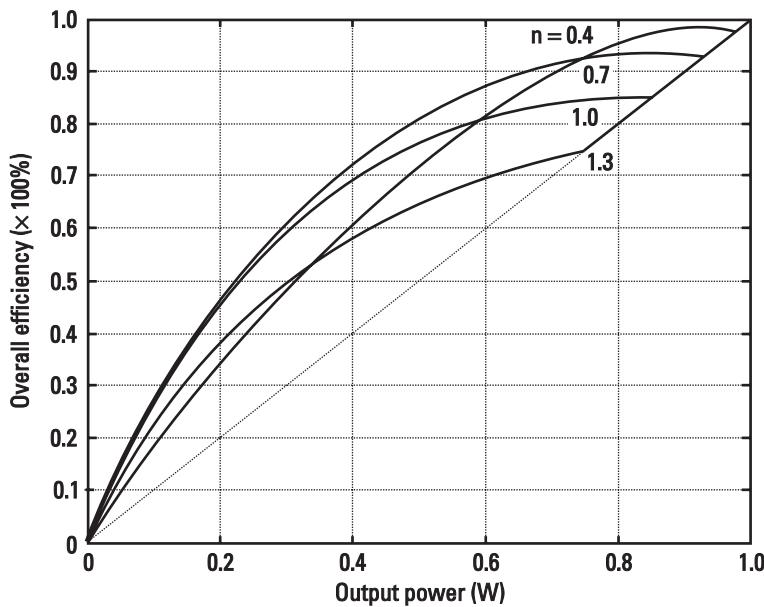


Figure 4.29 Overall efficiency as a function of the output power for various impedance transform ratios.

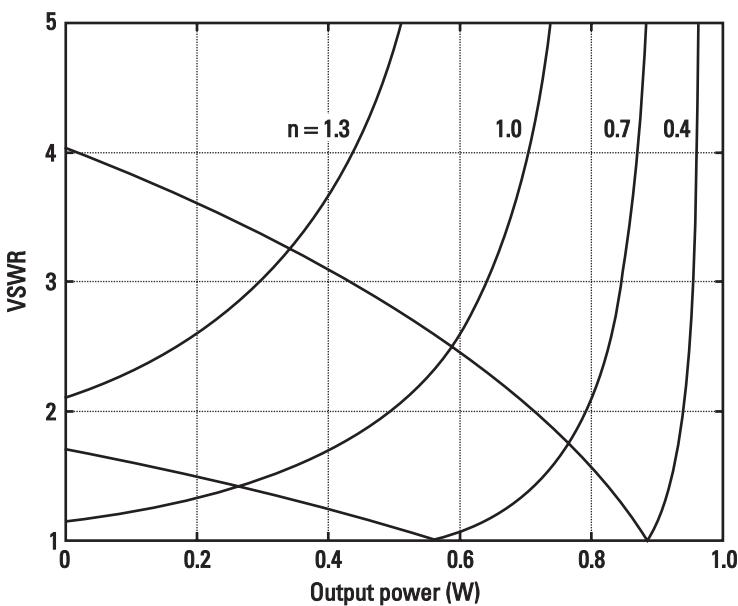


Figure 4.30 VSWR as a function of the output power for various impedance transform ratios.

recycling. On the other hand, when the transformation ratio “ n ” tends to zero, the diodes are most often turned on. In this case, the large signal impedance of the diodes approximates zero, and the largest portion of the RF power delivered to the recycling network will be reflected back to the isolator; thus the average power efficiency of the system is also low. Obviously, there exists an optimum impedance matching for the maximum average power efficiency, depending on the probability density of the modulated signal.

The probability density function of the modulated signal is required to calculate the power efficiency with and without the power recycling. The analytical expression of the probability density generally cannot be found, and hence the histogram from the simulation is used instead, as shown in Figure 4.16. The modulations under investigation include QPSK, OQPSK, $\pi/4$ -DQPSK, 16-QAM, and 64-QAM, with square-root raised cosine filtering. The roll-off factor of the shaping filter varies from 0.1 to 1.0, with an increase of 0.1. Figure 4.31 shows the calculated peak average power efficiency with the optimum impedance matching as a function of the roll-off factor for various modulations. The average power efficiency without

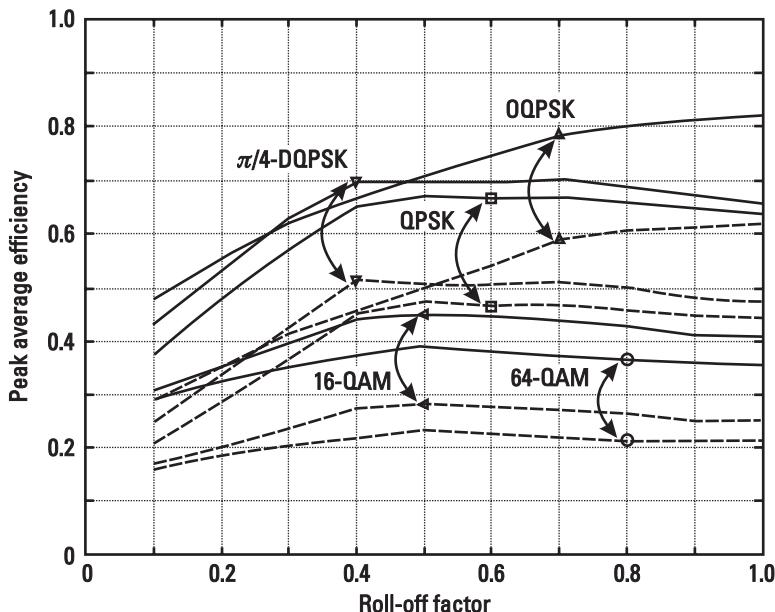


Figure 4.31 Peak average efficiency as a function of the roll-off factor for various modulations with and without recycling ($V_{\text{sup}} = 3V$, $R_d = 5\Omega$).

the recycling network is also displayed as the dashed lines to show the improvement. In this calculation, the maximum output power of the amplifier system is 1W. The supply voltage is 3V and the diode on-resistance is 5Ω . The recycling factor is hence 2.7, which corresponds to a maximum 57% recycling efficiency for the full-power level, and a maximum 65% recycling efficiency for the mid-power level, according to (4.120) and (4.121). As illustrated in Figure 4.31, with a properly chosen impedance-matching network, the net increase of the average power efficiency is between 14% and 21%, depending on the modulations. This corresponds to a relative improvement of the average power efficiency from 33% to 83%, which is quite respectable.

Note that the probability density functions in this calculation experiences quite different distributions, while the net increase of the average power efficiency is always approximately 16% for 16-QAM and 64-QAM, and approximately 20% for QPSK, OQPSK, and $\pi/4$ -DQPSK, except with the modest drop for the case of roll-off factor 0.1. As shown in Figure 4.16, the probability density of the $\pi/4$ -DQPSK-modulated signal typically possesses one or two peaks in certain medium-power levels and rolls off on both sides. The OQPSK-modulated signal roughly approximates a constant envelope; hence its probability density concentrates within a certain range of power levels well above dc. In the case of 64-QAM, the probability density function tends to be Gaussian and the PAP ratio is high. Very simply, the proper choice of R_d , V_{sup} , and P_{ava} is critical to the performance of the recycling network, and Figure 4.22 offers useful guidance for the estimation of the improvement.

The power-recycling experiment has been conducted for continuous-wave as well as linear-modulated signals [4, 5, 7]. The purpose of this experiment is to prove the concept of the power-recycling technique being a viable approach for gaining system efficiency. A prototype of the diode detector was tested with various power supply voltages, RF power levels, and modulation schemes.

The experiment for the continuous-wave signal was performed and the input power level was swept over several values and the supply voltage, was kept constant. Three different voltage values are used: 3, 4, and 5V, as shown in Figures 4.32 and 4.33. As can be seen from the results, the point at which the best match occurred did not coincide with the point for the best recycling efficiency. The maximum power transfer and optimum impedance are fairly close to each other, though the power resolution of the experiment could be improved. This result is consistent with the analysis shown in Figure 4.28 and was verified with the numerical calculation.

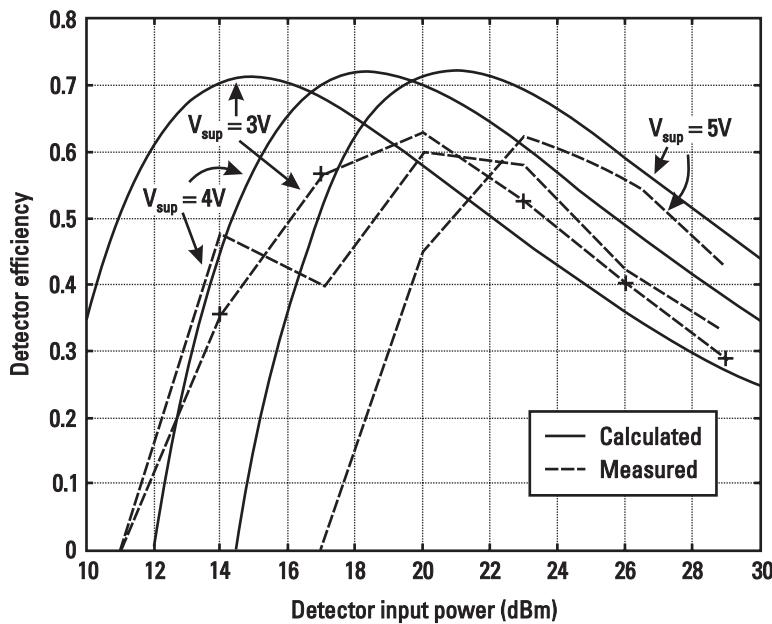


Figure 4.32 Comparison of calculated and measured recycling efficiency as a function of input power level [4].

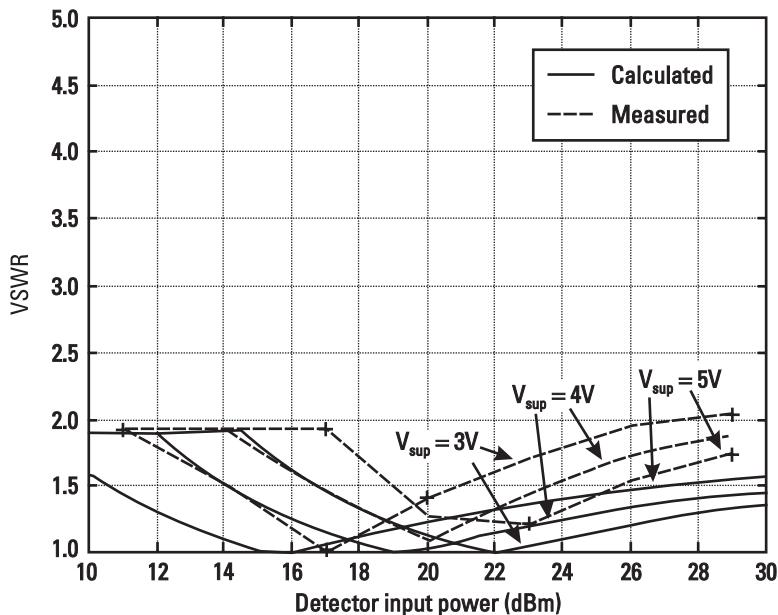


Figure 4.33 Comparison of calculated and measured variation of input VSWR as a function of input power level [4].

Figure 4.34 shows a plot of predicted amplifier system efficiency with and without the power-recycling network. For illustrative purposes, the power amplifiers in the system are assumed to have 100% efficiency. The recycling efficiency was varied between 0.9 and 0.1 to show the effects of power recycling on the overall power amplifier efficiency; the data from the detector circuit is also superimposed on the plot to compare it to the various contours of constant recycling efficiency. The “measured” data in Figure 4.34 refers to the measured recycling efficiency data of Figure 4.32; the system efficiency was then calculated based on that data. As can be seen, the overall power amplifier efficiency can be enhanced significantly by recycling the “wasted” power from the difference port of the 180° hybrid. The amount of improvement will depend on the modulation scheme and data-filtering technique applied.

The bandwidth of this technique is sufficiently broad to accommodate a variety of modulation schemes without significant degradation of the power returned to the power supply. Figure 4.35 plots the variation in efficiency using a 50% amplitude modulation of the input signal at a 10-kHz rate. In this case, there is little difference between this result and the

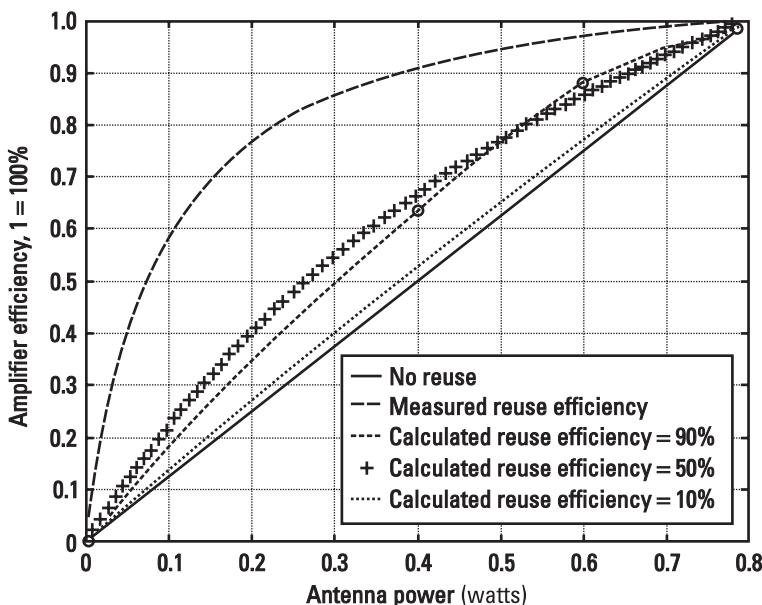


Figure 4.34 Calculated overall efficiency as a function of power delivered to the antenna for 0.8W peak outphasing power amplifier, showing the effect of recycling efficiency [4].

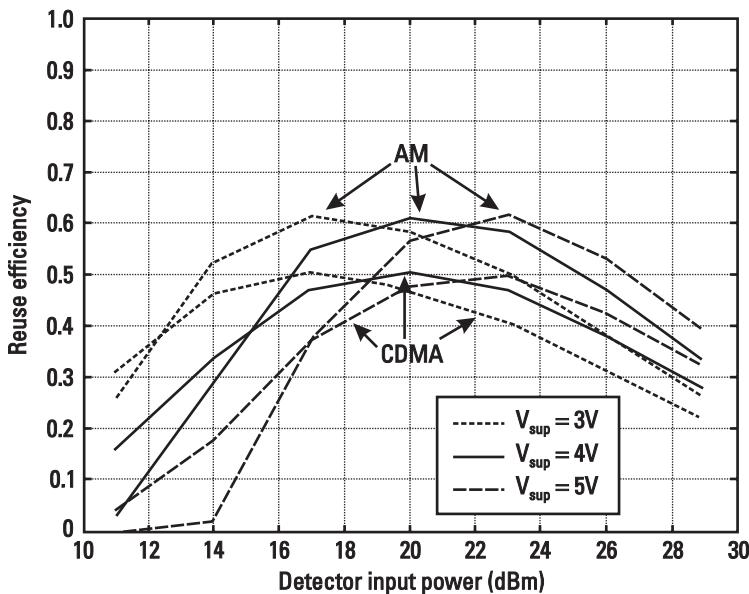


Figure 4.35 Measured variation of circuit efficiency as a function of input power level for 50% AM at 10 kHz and CDMA at 1.25 MHz at a carrier frequency of 1.96 GHz [4].

continuous-wave result of Figure 4.32. Even wider bandwidths are required for a CDMA-modulated signal, where the typical modulation bandwidth is increased to 1.25 MHz. In this case, the measured results show a loss in recycling efficiency of approximately 10%, yielding a peak recycling efficiency of approximately 50%.

4.9.3 Practical Implementations

For multicarrier channels in a base station, the probability density function of the modulated carrier tends to be Gaussian-distributed according to the central limit theorem. This will make the behavior of the recycling network much like that of a high-level modulation case, such as 64-QAM. The bandwidth of the power-recycling system is generally determined by the impedance-matching network. As demonstrated in [4], a simple design of the impedance-matching network provided sufficient bandwidth for a CDMA IS-95 modulated signal.

It is essential to reduce the diode loss to maximize the system performance. The diode “on-resistance” can be reduced by parallel connection

of a few diodes, but the shunt capacitance increases. The optimum design of the recycling network may be fine-tuned with SPICE simulations and experiments. It is also important to properly choose the matching network to prevent a high reverse voltage from being applied to the Schottky diodes to prevent breakdown.

It should be noted that the power-recycling circuit presented here is only one of several ways to approach this problem depending on the center frequency and the bandwidth of the signal. An alternative implementation of the power-recycling circuit at lower frequencies could be a transformer and diode pair, followed by a dc-dc converter. From a practical standpoint, a dc-dc converter would be desirable to provide the proper supply voltage to the power amplifiers in any given system.

References

- [1] Chireix, H., "High Power Outphasing Modulation," in *Proc. IRE*, Vol. 23, November 1935, pp. 1370–1392.
- [2] Cripps, S. C., *RF Power Amplifiers for Wireless Communications*, Norwood, MA: Artech House, 1999.
- [3] Sundstrom, L., and M. Johansson, "Effect of Modulation Scheme on LINC Transmitter Power Efficiency," *Elect. Lett.*, Vol. 30, No. 20, September 29, 1994, pp. 1643–1645.
- [4] Langridge, R., et al., "A Power Reuse Technique for Improved Efficiency of Outphasing Microwave Power Amplifiers," *IEEE Trans. Microwave Theory Technol.*, Vol. 47, No. 8, August 1999, pp. 1467–1470.
- [5] Zhang, X., et al., "Analysis of Power Recycling Techniques for RF and Microwave Outphasing Power Amplifiers," *IEEE Trans. cir. sys. II*, Vol. 49, No. 5, May 2002, pp. 312–320.
- [6] Pozar, D. M., *Microwave Engineering*, second edition, New York: John Wiley & Sons, Inc., 1998.
- [7] Langridge, R., "A Method for Enhancing the Efficiency of Power Amplifiers Used with Linear Modulation Schemes," Thesis (M.S.), University of California, San Diego, CA 1999.

Selected Bibliography

- Conradi, C. P., R. H. Johnston, and J. G. McRory, "Evaluation of a Lossless Combiner in a LINC Transmitter," *IEEE Canadian Conference on Electrical and Computer Engineering*, Edmonton, Alberta, Canada, May 1999, pp. 105–110.

- Cox, D. C., "Linear Amplification with Nonlinear Components," *IEEE Trans. Commun.*, Vol. COM-23, December 1974, pp. 1942–1945.
- Hammond, R., and J. Henry, "High Power Vector Summation Switching Power Amplifier Development," in *Proc. IEEE Power Electron. Specialists Conf. (PESC)*, Boulder, CO, June 1981, pp. 267–272.
- Hornak, T., and W. McFarland, "Vectorial Signal Combiner for Generating an Amplitude Modulated Carrier by Adding Two Phase Modulated Constant Envelope Carriers," U.S. Patent 5,345,189, September 1994.
- Krauss, H. L., C.W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, New York: John Wiley & Sons, Inc., 1980.
- Raab, F. H., "Efficiency of Outphasing RF Power-Amplifier Systems," *IEEE Trans. Commun.*, Vol. COM-33, No. 10, October 1985, pp. 1094–1099.
- Stengel, B., and W. R. Eisenstadt, "LINC Power Amplifier Combiner Method Efficiency Optimization," *IEEE Trans. Veh. Technol.*, Vol. 49, No. 1, January 2000, pp. 229–234.

Appendix 4A

4A.1 Available Power from the Hybrid Combiner

The diode conduction angle is determined by the power delivered to the recycling network, and therefore is related to the effective input impedance of the recycling network. On the other hand, the variation of the impedance alters the load presented to the power amplifiers, which in turn changes the power delivered to the recycling network. Hence, the input impedance and the available power to the recycling network are coupled together, depending on the practical situation. In Class A, AB, B, and C power amplifiers, the transistor is considered as a controlled current source. Variation of the load may cause transistor saturation or breakdown. In saturated Class AB, B, and C power amplifiers, the amplifiers can be regarded as a voltage source. For some switching-mode power amplifiers, like Class E, the amplifiers may fail to work if the load changes. Thus, the situation becomes rather complicated, and the maximum power transfer theorem generally does not apply to this situation. Let's imagine that an isolator is put in between the recycling network and the hybrid combiner. Now, looking back from the power amplifiers, the load is fixed; looking from the recycling network, the isolator acts like a voltage source and the maximum power transfer theorem applies.

Figure 4A.1 shows the block diagram of this approach. The power amplifiers are modeled as an ideal voltage source in series with a $50\text{-}\Omega$

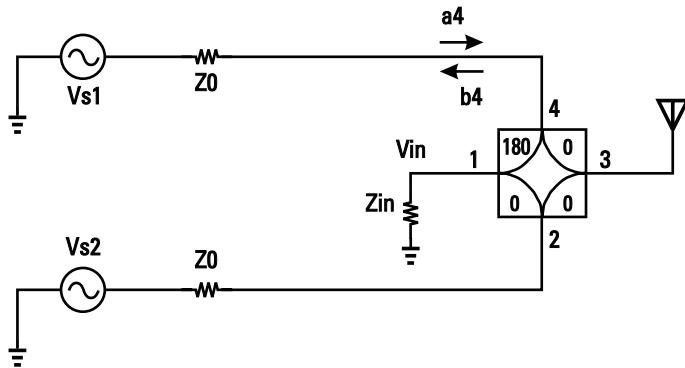


Figure 4A.1 Power available to the recycling network.

resistance. If the input impedance of the power-recycling network is Z_{in} , we have

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \quad (4A.1)$$

and it is easy to prove that

$$\begin{aligned} V_{s1} &= V_4 + Z_0 i_4 \\ &= 2\sqrt{Z_0} a_4 \end{aligned} \quad (4A.2)$$

and

$$\begin{aligned} V_{s2} &= V_2 + Z_0 i_2 \\ &= 2\sqrt{Z_0} a_2 \end{aligned} \quad (4A.3)$$

Thus, the voltage drop across the load is

$$\begin{aligned} V_{in} &= \sqrt{Z_0}(a_1 + b_1) \\ &= \sqrt{Z_0} \frac{2Z_{in}}{Z_{in} + Z_0} b_1 \\ &= \frac{1}{V_{s2} - V_{s1}} \frac{Z_{in}}{Z_{in} + Z_0} \end{aligned} \quad (4A.4)$$

Recall that $V_{s1} = s - e$ and $V_{s2} = s + e$; hence the power available to the recycling network is exactly the quadrature signal— $e(t)$ —portion of the total power.

4A.2 Recycling Efficiency and VSWR for Arbitrary Diode Model

The voltage-current characteristic of a Schottky barrier is usually described by empirical equations. More generally, the following arbitrary current-voltage relationship of the Schottky diode is assumed,

$$I_D = f(V_D) \quad (4A.5)$$

where V_D is the voltage drop across to the diode. Then, the dc component of the diode current can be found

$$i_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(V_{pk} \cos \phi - V_{sup}) d\phi \quad (4A.6)$$

and the fundamental component is

$$i_1 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(V_{pk} \cos \phi - V_{sup}) \cos \phi d\phi \quad (4A.7)$$

The same expressions for the available power (4.122) and the recycling efficiency (4.113) can also be obtained as a function of V_{pk} . The concept of diode conduction angle in this more general case is meaningless. The maximum recycling efficiency is determined by differentiating (4.113) with respect to the impedance transformation ratio “ n ” for the fixed available power to the recycling network. Similarly, (4.122) can be regarded as the implicit function of “ n ” and V_{pk} . Finally we have

$$\begin{aligned} \frac{d\eta_r}{dn} &= \frac{2V_{sup}}{P_{ava}} \frac{di}{dV_{pk}} \frac{dV_{pk}}{dn} \\ &= \frac{4V_{pk}}{P_{ava}} \frac{\int_{-\pi}^{\pi} f'(V_{pk} \cos \phi - V_{sup}) \cos \phi d\phi}{n\pi + \frac{Z_0}{n} \int_{-\pi}^{\pi} f'(V_{pk} \cos \phi - V_{sup}) \cos^2 \phi d\phi} \left(\frac{Z_0}{n^2} i_1 - V_{pk} \right) \end{aligned} \quad (4A.8)$$

The maximum recycling efficiency occurs when the above expression is equal to zero, which is equivalent to

$$\begin{aligned} Z_0 &= n^2 \frac{V_{\text{pk}}}{i_1} \\ &= Z_{\text{in}} \end{aligned} \tag{4A.9}$$

Equation (4A.9) concludes that in the case of the fixed available power to the recycling network, the maximum recycling efficiency and the lowest VSWR occur simultaneously. This conclusion is general and independent of the exact form of the diode model and is a direct consequence of the maximum power transfer theorem.

About the Authors

Xuejun Zhang received a B.S. in semiconductor physics from Peking University, China, in 1991, and an M.S. in electro-optics from the National University of Singapore, Singapore, in 1997. He received his Ph.D. in electrical engineering from the University of California, San Diego, in 2001. He was with the Southwestern Computing Center in China from 1991 to 1992 and with the Institute of Applied Electronics in China from 1992 to 1995. In these organizations he was involved in various research projects on laser optics, optical diagnostic systems, nonlinear optics, and optical materials, among others. Mr. Zhang is currently a senior RF engineer at Qualcomm Inc. in San Diego, California, where he is involved in integrated circuit design and research on the RF front end for wireless communications.

Lawrence E. Larson received a B.S. in electrical engineering in 1979, and an M. Eng. in 1980, both from Cornell University, Ithaca, New York. He received a Ph.D. in electrical engineering from the University of California, Los Angeles, in 1986.

Professor Larson joined the Hughes Research Laboratories in Malibu, California, in 1980, where he directed work on high-frequency InP, GaAs, and silicon-integrated circuit development for a variety of radar and communications applications. While at Hughes, he led the team that developed the first MEMS-based circuits for RF and microwave applications, the first InP-based HEMT foundry, and the development of the first Si/SiGe HBT microwave circuits with IBM. He was also the assistant program manager of the Hughes/DARPA MIMIC program from 1992 to 1994. From 1994 to 1996, he was at Hughes Network Systems in Germantown, Maryland, where he directed the development of radio frequency integrated circuits for wireless communications applications.

Professor Larson joined the faculty at the University of California, San Diego (USCD), in 1996, where he is the inaugural holder of the

communications industry chair. He has served as the director of the Center for Wireless Communications since July 2001. From July 2000 until July 2001, Professor Larson also served as the director of the IBM West Coast Design Center of Excellence in Encinitas, California (on leave of absence from UCSD), where he worked on the development of RFICs for 3G applications.

Among the awards received by Professor Larson are the Hughes Research Laboratories Sector Patent Award for his work on RF MEMS technology (1995); the 1996 Lawrence A. Hyland Patent Award of Hughes Electronics, for his work on low-noise millimeterwave HEMTs (corecipient); and the 1999 IBM Microelectronics General Managers Excellence Award for his work in Si/SiGe HBT technology 1993–1998 (corecipient).

Professor Larson was named one of the Wireless Influencers 2002, the top 200 people selected internationally as the top influencers in the wireless industry for 2002, and one of only 11 people selected from academia.

He has published more than 150 papers, has received 25 U.S. patents, and is a fellow of the IEEE.

Peter M. Asbeck is the Skyworks chair professor in the Department of Electrical and Computer Engineering at USCD. He attended MIT, where he received a B.S. and Ph.D. in 1969 and 1975, respectively, from the Electrical Engineering Department. He worked at the Sarnoff Research Center, Princeton, New Jersey, and at Philips Laboratory, Briarcliff Manor, New York, in the areas of quantum electronics and GaAlAs/GaAs laser physics and applications. In 1978, he joined the Rockwell International Science Center, where he was involved in the development of high-speed devices and circuits using III–V compounds and heterojunctions. He pioneered the effort to develop heterojunction bipolar transistors based on GaAlAs/GaAs and InAlAs/InGaAs materials, and has contributed widely in the areas of physics, fabrication, and applications of these devices.

In 1991, Dr. Asbeck joined the University of California at San Diego. His research interests are in the development of high-speed heterojunction transistors and their circuit applications.

Dr. Asbeck's research has led to more than 200 publications. He is a fellow of the IEEE and a distinguished lecturer of the IEEE Electron Device Society and the Microwave Theory and Techniques Society.

Index

- $\pi/4$ -DQPSK, 38–41
advantage, 38
combining efficiency as function of roll-off factor, 162
constellation and phase transitions, 40
defined, 38
disadvantage, 40
power probability density function, 161
- ACPR, 35
defined, 15–16
illustrated, 17
requirements characterization, 18
- Adaptive predistortion
defined, 23
illustrated, 24
See also Predistortion
- Adjacent channel interference (ACI), 18, 30
for bandlimited signals, 67
estimation with gain and phase mismatch, 64–67
output signal approximation, 65–66
in two-tone test, 63–64
- Advanced mobile phone system (AMPS), 1
- Alternating variable method, 103
- AM/AM conversion, 10, 76
- Amplitude-scaling factor, 61
- AM/PM conversion, 10, 22, 77
as consequence, 11
defined, 35
of nonlinear power amplifier, 104
- Analog-to-digital (A/D) converter, 91
bit length, 110
dc offset of, 97, 99
quantization error, 98, 110
word length, 111
- Application-specific DSPs (ASDSPs), 59, 75
- Arbitrary diode model, 188–89
- Arcsin, 55, 56
- Average noise power, 77
- Background calibration algorithm, 105–12
block diagram, 106
data transmission transparency, 111
defined, 105
implementation with RF switches, 111
measured lowpass-filtered signal, 113
practical considerations, 109–12
quadrature modulator error, 109
simulated lowpass-filtered signal during, 114
theory, 106–8
See also Calibration; Foreground calibration algorithm
- Baseband filtering, 41–50
- Baseband preconditioning, 89–90
block diagram, 90
defined, 89
- Baseband pulse shaping, 160
- Bessel filter, 81
- Bias adoption, 22
- Bit error rate (BER), 10
“Black box,” 151
- Butterworth filter, 80, 81
- Calibration
algorithm operations, 113
background, 105–12
foreground, 91–96
output spectra with/without, 115
- CDMA2000, 20

- Chireix, H., 27, 28
- Chireix power-combining, 142–45
 advantages, 159
 defined, 142–43
 efficiency, 143–44
 even-mode equivalent circuit, 150
 illustrates, 142
 linearity and, 145
 odd-mode equivalent circuit, 149
 summary, 145
See also Power combiners
- Class A amplifier
 design use, 136
 with lossy combining, 160
 transistor as current source, 135
- Class B amplifier
 current and voltage waveforms, 137
 design use, 136
 efficiency, 138
 efficiency optimization, 139
 efficiency reduction, 141
 efficiency variation, 141
 output power, 140
 output voltage, 137–38
 “overdriving,” 140
 schematic design, 137
 transistor as current source, 135
 voltage amplitudes, 139
- Class C amplifier
 design use, 136
 “overdriving,” 140
 transistor as current source, 135
- Class D amplifier, 135
 CMOS technology for, 147
 defined, 146
 equivalent circuit of, 148
 GaAs MESFET implementation, 157
 modeled as ideal square-wave voltage, 147
 MOS-based, with lossless combining, 147–56
 schematic, 146
 transistor power dissipation, 147
 voltage-mode, 136
- Class E amplifier, 135
 characteristics, 145
 schematic, 146
- Class F amplifier, 135, 136
- Code-division multiple access (CDMA), 4
 advantages, 4
 wideband (WCDMA), 20
See also IS-95 CDMA
- Combined analog locked-loop universal modulator (CALLUM), 27, 54, 88, 117, 119–24
 advantages, 117
 block diagram (CALLUM1), 120
 block diagram (CALLUM2), 121
 block diagram (CALLUM3), 123
 block diagram (CALLUM4), 124
 CALLUM1, 119–21
 CALLUM2, 121–23
 CALLUM3, 122–23
 CALLUM4, 123, 124
 disadvantage, 123
 dynamic behavior, 120
 feedback locked loop implementation, 119
 narrowband, 117
 sensitivity, 123
 structure illustration, 122
- Complementary cumulative distribution function (CCDF), 14, 15
- Complementary metal oxide-silicon (CMOS)
 implementation of Class D amplifier, 155
 inverters, 146
 SOS technology, 154, 155
 technology for Class D amplifiers, 147
- Correction schemes
 in background, 88
 based on training vectors, 88–101
- Differential binary PSK (DBPSK), 38
- Digital European cordless telephone (DECT), 89
- Digital modulation, 36–41
 $\pi/4$ -DQPSK, 38–41, 161, 162
 OQPSK, 4, 38, 67, 161, 162
 QAM, 36, 41, 87, 160, 162
 QPSK, 36–41, 161, 162
- Digital-to-analog (D/A) converters, 79, 80
- Dirac function, 43

- Direct search algorithm, 104–5
baseband block diagram, 105
defined, 104
- Distortion
AM/AM, 22, 23
AM/PM, 23
harmonic, 9–10
intermodulation, 9, 11
- Doherty amplifier, 22
- Double-sideband suppressed carrier (DSB-SC) signals, 98
- DSPs
application-specific (ASDSPs), 59, 75
baseband, 91, 107
sampling rate, 79–81
- Envelope elimination and restoration (EER), 22
- Error vector magnitude (EVM)
measurement, 35
defined, 18
illustrated, 19
as root-mean-square (rms) value, 18
- Feedback
linearization with, 25
signal component separation with, 57
- Feedback loop
inverse sine phase modulator with, 55
signal component separator with, 57
VLLs, 126
- Feedforward
predistortion illustration, 27
techniques, 26
- Field effect transistors (FETs), 136, 156
- Filters
Bessel, 81
Butterworth, 80, 81
FIR, 113
Gaussian, 49–50
IS-95 baseband, 50
Nyquist, 45, 47
raised cosine, 45–49
reconstruction, 79–81
two-stage polyphase, 71
- Finite impulse response (FIR) filters, 113
- Foreground calibration algorithm, 91–96
generation and measurement, 92
measured output spectra with/without, 101
path imbalance/calibration signals
relationship, 94
practical considerations, 96–101
signal waveforms, 100
See also Background calibration algorithm; Calibration
- Frequency-division multiple access (FDMA), 2–3
- Frequency shift keying (FSK), 3
- Gain imbalance, 89, 96
background calibration, 109
compensation, 90
estimation error, 99
- Gaussian filter, 49–50
defined, 49
frequency response, 49
high-frequency components and, 49–50
impulse response, 49
- Gaussian minimum-frequency shift keying (GMSK), 3, 49
- Global system for mobile communication (GSM), 3, 49, 50
- Havens quadrature signal generator, 71
- Hybrid combiners
available power from, 186–88
outphasing amplifier with, 158
transformer, 131
See also Power combiners
- Impedance transform ratio, 173
- Input intercept point (IIP), 13
- Input signal amplitude, 13
- Intermodulation distortion, 11
illustrated, 9
output frequency generation, 11
two-tone test characterization, 11
- Intermodulation frequencies, 9
- Intermodulation products
second-order, 11, 12
third-order, 11–12
- Intersymbol interference (ISI), 43
due to lowpass filtering, 44
elimination, 45

- IS-95 CDMA, 4
 - baseband filter, 50
 - simulated spectrum, 67
 - simulated total output spectrum, 82
 - See also* Code-division multiple access (CDMA)
- Japanese TACS/narrowband TACS (JTACS/NTACS), 2
- Least square (LS) algorithm, 113
- Linearity
 - degradation, 35–36, 73, 81
 - effects of reconstruction filter and DSP sampling rate, 79–81
 - performance, 35–83
 - quadrature modulator error effects on, 67–75
 - two-tone test analysis, 62–64
- Linearization
 - with feedback, 25
 - techniques, *xi*
- Linear-modulated signals, 176–84
 - average power efficiency, 178
 - average power efficiency without power recycling, 180–81
 - instantaneous overall efficiency, 176, 178
 - overall efficiency as function of output power, 179
 - overall power efficiency, 176
 - peak average efficiency, 180
 - probability density function, 180
- Local oscillator (LO), 68, 69
- Lookup table (LUT), 49, 59
 - one-dimensional, 61
 - symmetry, 60
- Lossless power combiners, 133, 134, 135
- Lossy power combiners, 130
 - application to outphasing power amplifiers, 156–59
 - instantaneous combining efficiency, 159
 - wasted power, 159
 - See also* Power combiners
- Lowpass filter (LPF), 96, 97
- Matched, isolated combiners, 133, 134
- Memoryless weak nonlinearities, 7
- Metal semiconductor FET (MESFET) technology, 156
- Mismatch correction scheme, 112–14
- Modulation. *See* Digital modulation
- MOS-based Class D amplifier, 147–56
 - analysis, 147–53
 - calculated power efficiency, 154
 - fundamental component, 151
 - fundamental component (even-mode operation), 152–53
 - implementation, 155
 - output power, 151
 - parasitic capacitor, 154
 - simulation of, 153–56
 - source current, 152
 - system efficiency, 153
 - voltage generators, 152
 - See also* Class D amplifier
- Newton-Raphson algorithm, 71
- Nonlinear inverse sine phase modulator, 52
- Nonlinearities, 83
 - illustrated, 7
 - memoryless weak, 7
 - strong, 13
- Nyquist filters, 45, 47
- Nyquist's first method, 44–45
- Offset quadrature PSK (OQPSK), 4, 38, 67
 - combining efficiency as function of roll-off factor, 162
 - constellation and phase transitions, 39
 - power probability density function, 161
- Q-channel data stream, 38
 - use, 38
 - See also* Digital modulation
- Organization, this book, *xii–xiii*
- Orthogonal frequency division multiplexing (OFDM), 112
- Out-of-band interference, 35
- Outphasing approach
 - concept, 27
 - defined, *xi*
 - theory, 29–32
- Outphasing power amplifiers, 27–32
 - amplifier choices, 135–36
 - band-limited input, 64

- with Class A, B, and C amplifiers, 137–42
concept description, 29
with foreground calibration loop, 91
historical perspectives, 27–29
with hybrid power combining, 158
linearity performance, 35–83
lossy power combiners application, 156–59
path mismatch reduction techniques, 87–126
with phase imbalance correction, 102
power-combining techniques, 130–35
power recycling, 163–85
quadrature modulator errors and, 72–75
SCS quantization error effects on, 75–79
signal component separation, 51–61
with simplex search algorithm, 103
simplified block diagram, 28
with transmission line coupling, 32
Output intercept point (OIP), 13
Output power
Class B amplifier, 140
MOS-based Class D amplifier, 151
probability, 160
probability distribution, 159–63
- Path imbalance, 61–67, 81–82
ACI estimation and, 64–67
correction, 61
foreground calibration path convergence and, 95
manifestation, 61
two-tone test and, 62–64
- Path mismatch errors
baseband preconditioning of, 89–90
foreground calibration algorithm of, 91–101
- Path mismatch reduction techniques, 87–126
background calibration algorithm, 105–12
for broadband applications, 112–14
direct search algorithm, 104–5
introduction, 87–88
phase-only correction, 101–
- simplex search algorithm correction, 102–4
training vector-based, 88–101
transparent to data transmission, 101–12
VCO-derived synthesis, 114–26
- Peak-to-average power (PAP) ratio, 14
defined, 14
inverse, 76
probabilistic measurement of, 15
- Phase imbalance, 95
background calibration, 109
compensation, 90
determination, 108
estimation error, 99
- Phase-locked loops (PLLs), 114
block diagram, 117
cross-coupled, 115
defined, 119
- Phase-only correction, 101–2
defined, 101
outphasing amplifier with, 102
- Power-added efficiency (PAE), 20
- Power amplifiers
average efficiency, 159
characterization for wireless communications, 4–22
large-signal nonlinear, 14
linearization and efficiency-enhancement techniques, 22–27
nonlinear distortion, 6
power efficiency measurements, 19–22
role in wireless communication systems, 1–4
switching-mode, 79
waveform quality measurements, 6–19
See also Outphasing power amplifiers
- Power combiners, 130–35
admittance matrix, 134
behavior, 133
Chireix, 142–45
efficiency, 129
hybrid, 131, 158, 186–88
impedance matrix, 134
as linear circuits, 130
lossless, 133, 134, 135
lossy, 130, 156–59

- Power combiners (continued)
 matched, isolated, 133, 134
 for minimizing interactions between
 input ports, 132
 “Rat-Race,” 131
 reactive, 132
 S matrix, 130, 132, 133
 structures, 131
 transform-based, 132
 transmission line-based, 132, 135
 Wilkinson, 131, 157
- Power efficiency
 lower results, 22
 measurements, 19–22
 PAE, 20
 poor, origins of, 21
- Power probability density function, 161
- Power recycling, 163–85
 alternative implementation, 185
 analysis for continuous-wave signal,
 164–76
 available power, 168
 average power efficiency without,
 180–81
 bandwidth, 184
 defined, 163
 diode voltage, 165
 efficiency, 167, 168–69, 171, 172, 174
 factor, 172, 173
 fundamental component, 165–66
 illustrated, 163
 impedance transform ratio,
 166, 173, 174
 optimum design, 185
 practical implementation, 184–85
 reflection coefficient, 167
 RF-dc power conversion, 166
 Schottky diodes in, 163
 series inductor, 166
 source voltage, 168
 VSWR, 167
- Power series expansion, 8
- Predistortion
 adaptive, 23, 24
 concept illustration, 23
 defined, 22–23
 feed-forward, 27
- practical limitations, 24
 problems, 26
- P-type MOS (PMOS), 155
- Quadrature amplitude modulation (QAM),
 36
 16-QAM, 42, 160
 64-QAM, 160
 combining efficiency as function of
 roll-off factor, 162
 defined, 41
 M-QAM, 87
 power probability density function, 161
See also Digital modulation
- Quadrature modulator errors, 67–75
 adjacent channel residue, 96
 background calibration, 109
 compensation, 89
 correction by test vectors, 72
 effects on linearity, 67–75
 effects on outphasing systems, 72–75
 generation, 69
 minimization, 69–72
 presence of, 96
 setting upper limit, 97
- Quadrature modulators
 block diagram, 68
 constant envelope characteristics and,
 73
 gain and phase mismatch, 73, 74
 perfectly balanced, 96
- Quadrature PSK (QPSK), 36–41
 $\pi/4$ -DQPSK, 38–41, 161, 162
 combining efficiency as function of
 roll-off factor, 162
 defined, 36–37
 offset (OQPSK), 4, 38, 67, 161, 162
 power probability density function, 161
 realization, 37
 signal constellation and phase
 transitions, 37
 waveforms, 38
See also Digital modulation
- Quadrature signal
 calculation, 53
 generation, 56
 generation with RC-CR network, 70

- Havens generator, 71
in-phase component, 55
phase difference, 56
quantization error effects of, 76–79
wideband, 65
- Quantization error effects, 75–79
A/D converter, 98, 110
of quadrature signal, 76–79
SNR, 78
of source signal, 75–76
- Quantization noise, 78, 82
- Raised cosine filter, 45–49
cutoff frequency, 47
defined, 45
frequency response, 45–46
implementation, 48
impulse response, 46–47, 48
roll-off factors, 46
square-root, 48
- “Rat-Race” combiner, 131
- RC-Cr network, quadrature signal generation with, 70
- Reactive combiners, 132
- Reconstruction filter, 79–81
- Recycling efficiency
for arbitrary diode model, 188–89
calculated and measured comparison, 182
calculated overall, 183
calculating, 168
as function of available power, 176
as function of impedance transform ratio, 173
as function of “ n ,” 169
maximum, 171, 172, 174, 178
measured variation of, 184
plots, 169, 170
See also Power recycling
- Recycling factor, 172, 173
- RF-dc power conversion, 166
- Schottky barrier, 188
- Schottky diodes, 163, 188
- SCS, 52
analog implementation, 31
imbalanced caused by, 36
in-phase/quadrature, 54
- quantization, 36
quantization error effects, 75–79
- Second-order intermodulation
output, 12
products, 11
- Shannon’s channel capacity formula, 42
- Signal component separation, 51–61
canonical forms, 51
with feedback loop, 57
with frequency translation, 59
with in-phase/quadrature method, 54
with nonlinear phase modulators, 51
with power feedback, 57
- Signals
accuracy, 18
bandpass, separation of, 29, 30
constant envelope, 29
DSB-SC, 98
linear-modulated, 176–84
quadrature, 53, 55, 56
spectral growth, 18
- Signal-to-noise ratio (SNR), 42
in calibration loop, 108
due to quantization signal error, 78
- Silicon on sapphire (SOS) CMOS
implementation, 155
technology, 154
- Simplex search algorithm, 102–4
alternating variable method, 103
block diagram, 103
defined, 102
with two variables, 104
- S matrix, 130, 132, 133
- Taylor expansion, 95
- Third-order intermodulation
fractional, 11–12
products, 11
- Time-division duplex (TDD), 89
- Time-division multiple access (TDMA), 3
- Total access communications system
(TACS), 2
- Transform-based combiners, 132
- Transmission line-based combiners, 132, 135
- Traveling wave tube amplifiers (TWTAs), 26

- Two-stage polyphase filter, 71
- Two-tone tests, 7
- acceptance, 62
 - linearity analysis, 62–64
 - output spectrum, 64
 - performance, 12
- Variable gain amplifier (VGA), 56
- control signal, 57
 - control voltage, 58
 - gain, 56, 57
- VCO-derived synthesis, 114–26
- Vector-locked loops (VLLs), 54, 88, 124–26
- advantage, 117
 - control equations, 125
 - disadvantage, 126
 - feedback loop, 126
 - generic control equations, 124
 - illustrated, 125
 - narrowband, 117
- Very large-scale integration (VLSI)
- technology, *xi*
- Voltage-controlled oscillators (VCOs)
- combined outputs, 116
 - control voltages, 118, 119
 - free-running frequency, 118
 - as integrator, 118
- output, 115
- output phase, 118
- output signals, 118
- sensitivity, 118
- Voltage standing-wave ratio (VSWR), 26
- for arbitrary diode model, 188–89
 - calculated and measured comparison, 182
- calculating, on recycling network, 164
- as function of available power, 177
 - as function of impedance transform ratio, 173, 175
 - as function of output power, 179
 - lowest, 169, 175
 - optimum, 171
 - recycling network, 167
 - under optimum efficiency, 175
- Volterra-series analysis, 10
- Waveform quality measurements, 6–19
- Wideband CDMA (WCDMA), 20
- Wilkinson combiner, 131, 157
- Wireless systems
- 1G, 2
 - 2G, 5
 - power amplifier characterization for, 4–22
 - power amplifier role in, 1–4