UNIVERSITY OF CALIFORNIA
RIVERSIDE

High Speed ADC Design Methodology

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

He Tang

December 2010

Dissertation Committee:

Prof. Albert Wang, Chairperson
Prof. Roger Lake
Prof. Sheldon Tan

The Dissertation of He Tang is approved:

_____

_____

_____
Committee Chairperson

University of California, Riverside

Acknowledgements

First and foremost, I am heartily thankful to my advisor, Prof. Albert Wang, whose encouragement, guidance and support from the beginning to the end has enabled me to develop comprehensive research skills. My academic achievement would never have been possible without his supervision. He pointed me into the right direction for research. Not only his dedication to research work, but also his integrity and personality in everyday life have always encouraged me throughout my Ph.D. studies.

My deepest love and gratitude go to my mother and my father. They have provided their greatest supports to me. I can hardly find any word that could express my gratitude and love to them. I wish my parents healthy and happy.

It is an honor for me to have three distinguished professors, Prof. Roger Lake, Prof. Sheldon Tan and Prof. Albert Wang, to be my Ph.D. committee members. I also thank my good lab mates, Xin Wang, Hui Zhao, Lin Lin, Qiang Fang, Siqiang Fan, Zitao Shi and my other friends who gave me so much helps and encouragement, and filled my life with joy. Special thanks would be given to Hui Zhao and Siqiang Fan who have helped me a lot on my research work, and Qiang Fang and Haipeng Ding for sharing the gym time with me every day!

I appreciate OmniVision Technologies Inc. and SMIC Corp. for supporting my internship and research design.

Lastly, I offer my regards and blessings to all of those who supported me in any respect during my Ph.D. studies.

To my parents for all the support.

In memory of my grandfather.

ABSTRACT OF THE DISSERTATION

High Speed ADC Design Methodology

by

He Tang

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, December, 2010
Prof. Albert Wang, Chairperson

Analog-to-digital converter (ADC) is a very fundamental and key part to nearly all kinds of electronics. The applications cover a wide range requiring different resolution to different sampling rate, including UWB systems, radar detection, wide band radio receivers, optical communication links, CCD imaging, ultrasonic medical imaging, digital receivers, base stations, digital video (for example, HDTV), xDSL, cable modems, and fast Ethernet. Among them, lower resolution very high speed ADC is a critical part for building UWB system, disk drive read channels and optical communication.

This thesis consists of two parts. The first part focuses on the design of a high speed low resolution flash ADC in 90nm technology. Capacitive interpolation technique was used in this flash ADC in order to reduce the hardware requirement and input capacitance. No sample-and-hold (S/H) circuit is needed since the distributed capacitors

(including capacitors in the very front end and the interpolated capacitors) serve to sample and hold the signals. Offset cancellation and averaging techniques are also implemented to reduce the offsets and the non-linearity. The ADC design achieves a sampling speed of 2.3GSps with 4 bits resolution in 90nm CMOS technology.

The second part describes a new comprehensive ADC design methodology for capacitive interpolated flash ADCs, aiming to provide a quantitative, yet handy design guideline for circuit designers to conduct practical ADC designs. This new ADC design methodology provides a quantitative and comprehensive mapping between ADC chip level performance specs and various design parameters at different levels, such as, interpolation factor, number of stages, pre-amplifier bandwidth, loading effects, transistor sizes, technology parameters and etc. It serves to allow IC designers to conduct quick and quantitative flash ADC designs for well-balanced overall chip performance in practices. A dynamic power consumption analysis technique for capacitive interpolated flash ADCs is also discussed.

Index terms: flash ADC, high speed, interpolation, design methodology, conversion rate, resolution, dynamic power dissipation.

Contents

List of Figures

List of Tables

List of Symbols and Abbreviations

Abbreviations	Definition
ADC	Analog-to-Digital Converter
BER	Bit Error Rate
BGR	Band-gap Reference
CMFB	Common-Mode Feedback
DAC	Digital-to-Analog Converter
DNL	Differential Non-Linearity
DR	Dynamic Range
DSP	Digital Signal Processing
ENOB	Effective Number Of Bits
FS	Full Scale
IC	Integrated Circuit
IF	Intermediate Frequency
INL	Integral Non-Linearity
LSB	Least Significant Bit
MDAC	Multiplying Digital-to-Analog Converter
MSB	Most Significant Bit
RF	Radio Frequency
SAR	Successive Approximation Register
SC	Switched Capacitor
SFDR	Spurious-Free Dynamic Range
SNDR	Signal-to-Noise-and-Distortion Ratio
SNR	Signal-to-Noise Ratio

SOC	System-On-a-Chip
TH	Track and Hold
THD	Total Harmonic Distortion

CHAPTER 1

INTRODUCTION

1.1 Motivation of ADC

Analog-to-digital converter (ADC) is mostly commonly used in nearly all kinds of electronics. In real world, signals such as voice, image and other information are analog, but in electronic devices, only digital signals can be processed, which means all analog signals needs to be converted into digital signals. An ADC is supposed to complete such conversions. It converts the analog signals (voltages, currents and etc.) into digital signals (normally binary), which will be processed by a DSP in electronic devices. On the other hand, a digital-to-analog converter (DAC) performs the opposite way (DAC is not discussed in this thesis). Digital signals processed by a DSP are sent to a DAC and converted to analog signals so that people can hear the music, see the image, and etc.

Therefore, it is obvious that the AD/DA converter is an indispensable part and plays a key role. It is like a translator connecting the real world and the electronic devices. Whatever the electronics are evolving, since the real world is analog, the AD/DA converter can never be disappeared. This may be the only motivation of researching and designing AD/DA converters. Figure 1.1 shows how this AD/DA conversion process works.

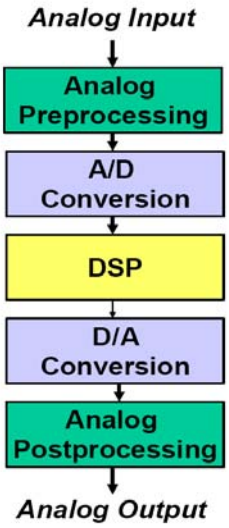


Figure 1.1 AD/DA Conversion Process.

There are many different types of AD converters which are used for quite distinct purposes. People design ADCs according to their specification requirements. Among all kinds of ADCs, high speed ADC is becoming more and more important and widely applied nowadays. For example, low-to-medium bit resolution very high speed (4-8 bits and over several GSps speed) ADCs have applications in UWB systems, disk driver, radar detection, wide band radio receivers and optical communication links, while medium-to-high bit resolution high speed ADCs (8-14 bits and hundreds of MSps to 1-2 GSps speed) have application in CCD imaging, ultrasonic medical imaging, digital receivers, base stations, digital video (for example, HDTV), xDSL, cable modems, and fast Ethernet.

This research mainly focuses on low resolution very high speed flash ADC (4-bit 2GSps capacitive interpolated flash ADC) design and its methodology. The design

methodology is novel and creative work which has never been researched before, and finally exciting results are obtained. This new methodology builds up a quantitative mapping between ADC performance (sampling rate and resolution bits) and design parameters, such as, interpolation factor, number of stages, pre-amplifier bandwidth, loading effects, transistor size, etc. Thus it provides an accurate and quick bottom-up design method for capacitive interpolated flash ADC design and its optimization

1.2 Principles of ADC Operation

The principle of how an ADC is working is shown in Figure 1.2 [1]. The continuous-time input signal is first sampled by a sample-and-hold (S/H) circuit and transformed into a discrete-time signal. Then the sampled signal is quantized and given a quantization level approximately. The last step is encoding, each quantization level encoded into a binary number. This binary number is the final output for an ADC and will be sent to the next stage – DSP.

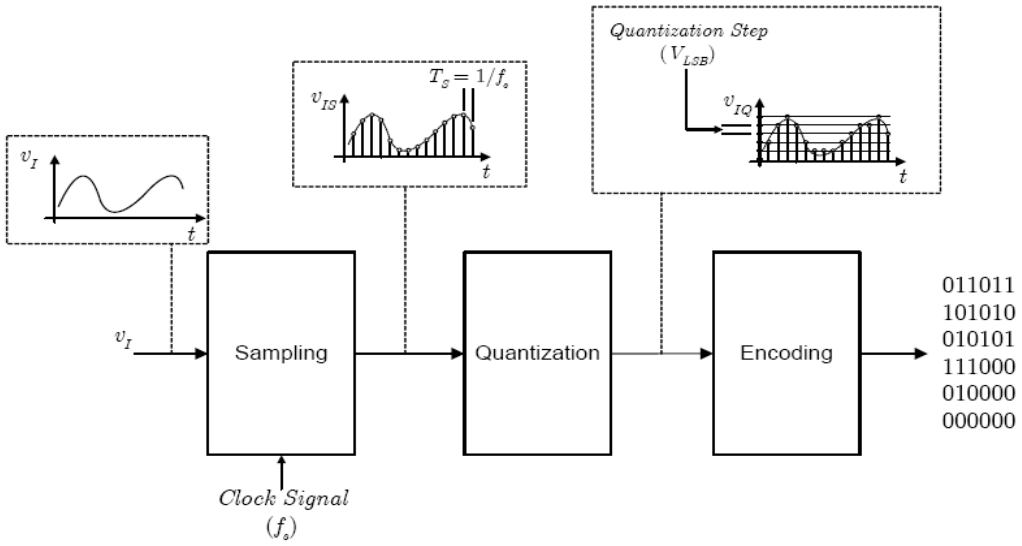


Figure 1.2 The Principle Operation of an ADC.

During S/H operation, the input signal is sampled with a sampling frequency of f_s, which means the S/H circuit is tracking the input signal in the first phase and held for a certain value in the second phase. According to Nyquist theorem, as long as the bandwidth of the input signal is less than $f_s/2$, the input signal can be recovered without any information lost [1].

In the quantization, the full scale of the input range is divided into 2^N steps by a quantizer, where each step is equal to $V_{LSB}=V_{FS}/2^N$. Since the sampled signals will be mapped to the discrete quantization levels, it is obvious that irreversible errors, quantization errors, are introduces, which prevents the exact the reconstruction [1]. The transfer function of quantization errors is shown in Figure 1.3 [1]. From the Figure, the maximum quantization error is easily obtained as $|e_q| \le V_{LSB}/2$.

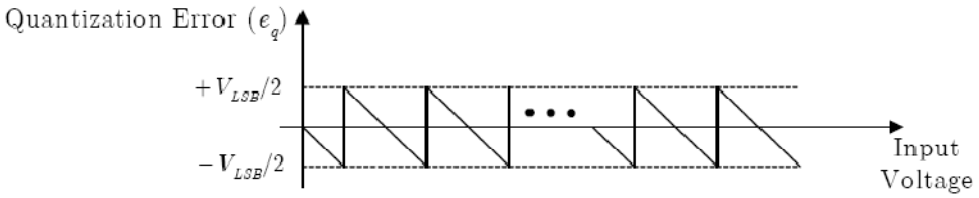


Figure 1.3 The Transfer Function of Quantization Errors.

The binary outputs are encoded by an encoder. Normally, for suppressing sparkle code and metastability, a string of thermomteter codes (outputs of the quantization block) are first encoded into Grey codes, and then binary outputs. The overall transfer function of an ADC is illustrated in Figure 1.4 [1].

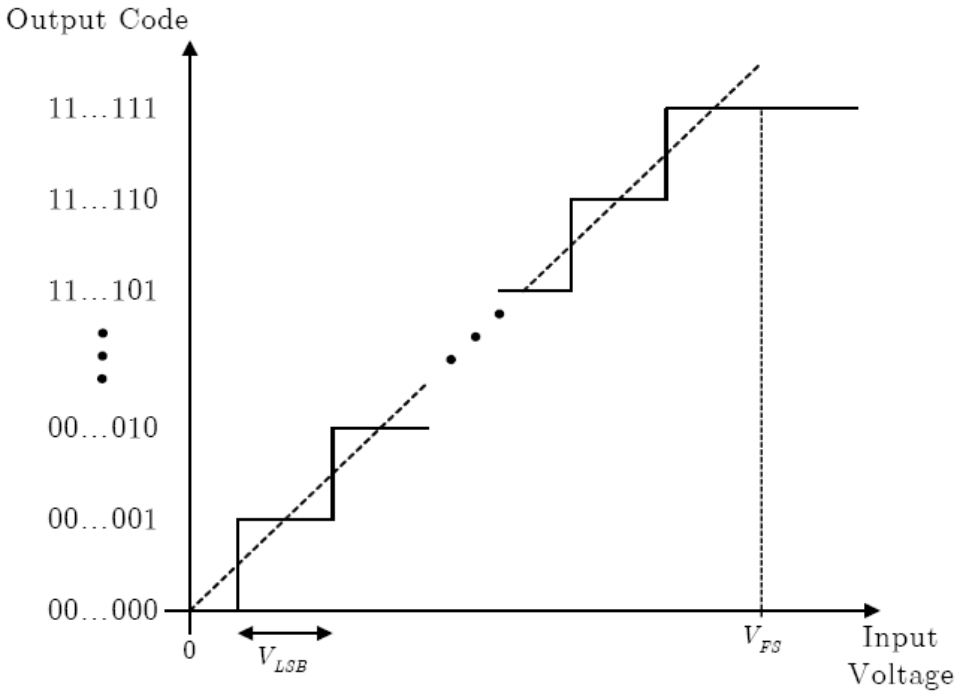


Figure 1.4 The Transfer Function of an ideal ADC.

1.3 Characterization of ADC

Analog to digital converters are characterized in a number of different ways. Some of them are measured as static performance, and some are measured in frequency domain.

1.3.1 Static Performance

1.3.1.1 Differential Nonlinearity

The step size in the non-ideal data converter deviates from the ideal size Δ and this error is called the differential nonlinearity (DNL) error [1]. Simply speaking, it is the measure of the maximum deviation from the ideal step size of 1 LSB.

For a DAC the DNL can be defined as the difference between two adjacent analog outputs minus the ideal step size; while for the ADC, it is the difference between transition points. The normalized expression is:

$$\begin{aligned} DNL_k &= \frac{\tilde{x}_{t,k+1} - \tilde{x}_{t,k} - \Delta}{\Delta} \ (ADC) \\ &= \frac{\tilde{x}_{a,k+1} - \tilde{x}_{a,k} - \Delta}{\Delta} \ (DAC) \end{aligned} \tag{1.1}$$

where $\tilde{x}_{t,k+1}$ and $\tilde{x}_{t,k}$ are the analog inputs for AD converters, and $\tilde{x}_{a,k+1}$ and $\tilde{x}_{a,k}$ are the analog outputs for DA converters.

Figure 1.5 shows the DNL errors below [2]:

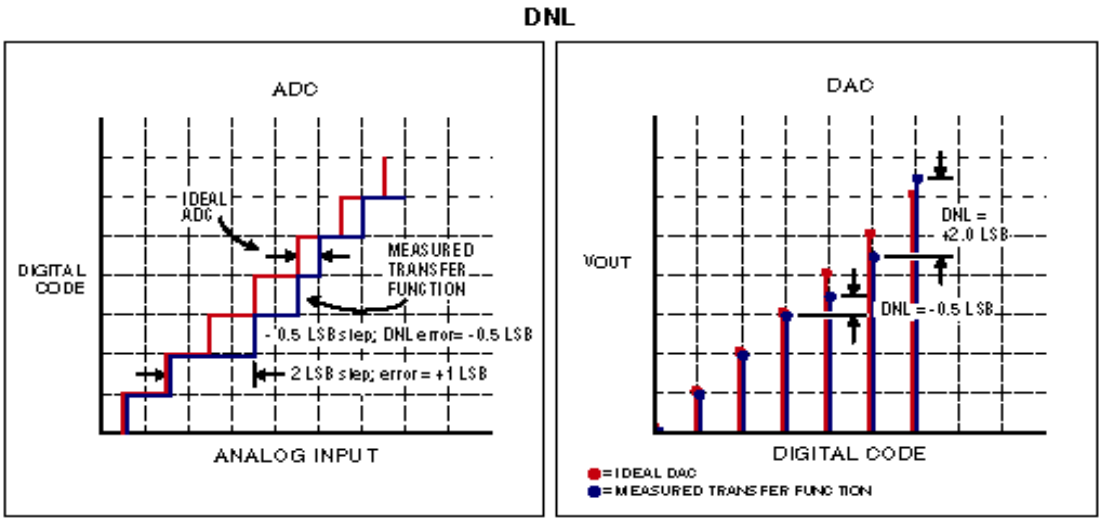


Figure 1.5 DNL Errors in ADC and DAC.

1.3.1.2 Integral Nonlinearity

The total deviation of an analog value from the ideal value is called integral nonlinearity (INL). It is the deviation of the entire transfer function from the ideal function. The normalized INL is expressed as [1]:

$$\begin{aligned} INL_k &= \frac{\tilde{x}_{a,k} - x_{a,k}}{\Delta} \\ &= \sum_{i=1}^{k-1} DNL(i) \qquad (\text{For both } ADC \text{ and } DAC) \end{aligned} \tag{1.2}$$

where $\tilde{x}_{a,k}$ and $x_{a,k}$ are the actual and ideal outputs of the converter respectively.

Figure 1.6 shows the INL errors below [2]:

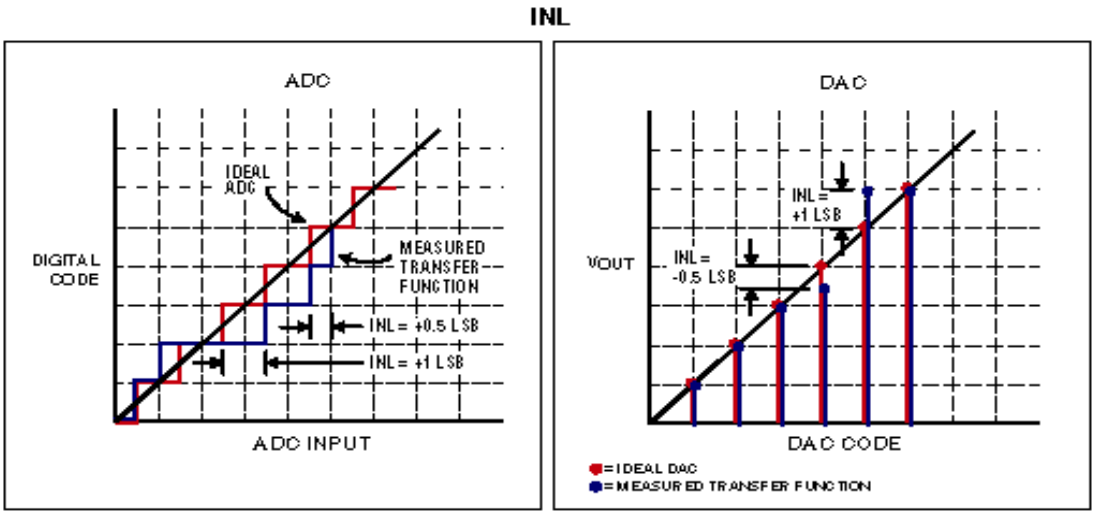


Figure 1.6 INL Errors in ADC and DAC.

1.3.1.3 Offset Errors

The offset error is difference between the ideal LSB transition point to the actual transition point. It is calculated as [1]:

$$X_{offset} = \frac{1}{2^N} \sum_{k=0}^{2^N-1} (\tilde{x}_{a,k} - x_{a,k}) \tag{1.3}$$

Figure 1.7 shows the offset errors below [2]:

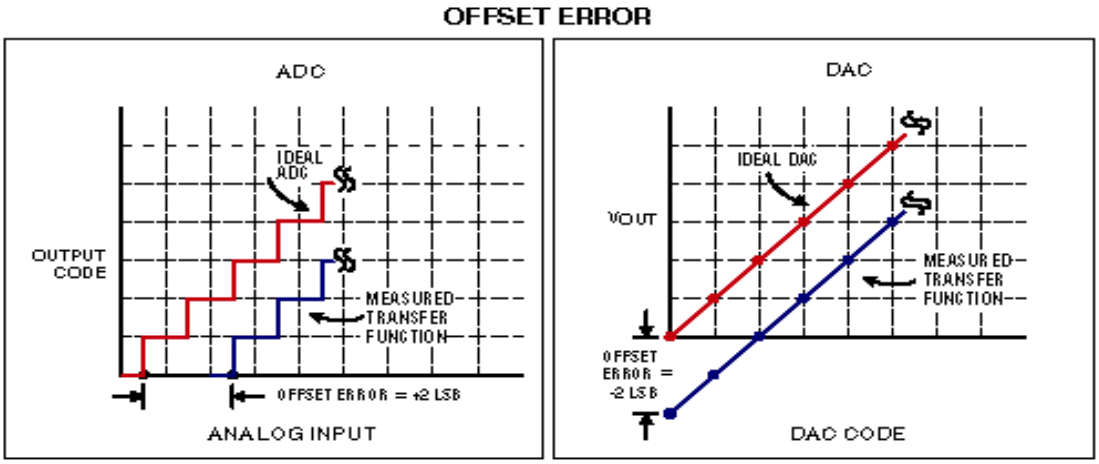


Figure 1.7 Offset Errors in ADC and DAC.

1.3.1.4 Gain Errors

The gain error is the deviation from the slope of actual transfer function to the slope of ideal transfer function. It is calculated as [1]:

$$\tilde{X}_a = AX_a + X_{offset}$$
$$A = \frac{\left\langle \tilde{X}_{a,k} \cdot X_{a,k} \right\rangle}{\left\langle X_{a,k}^2 \right\rangle} \tag{1.4}$$

Figure 1.8 shows the gain errors below [2]:

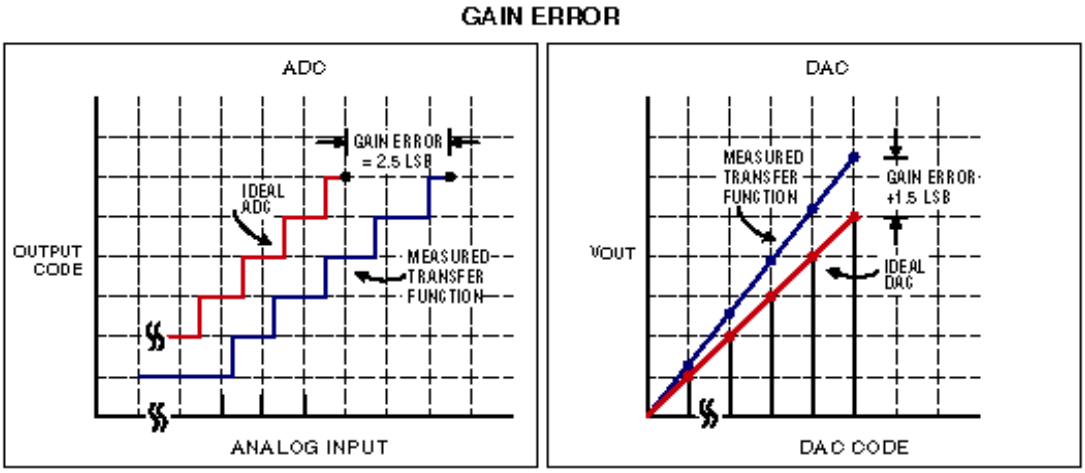


Figure 1.8 Gain Errors in ADC and DAC.

1.3.2 Frequency Domain Measures

For data converters, the static performances (DNL, INL, offset and gain errors) are not sufficient to characterize the ADC. It is more convenient to characterize the ADC in frequency domain by measuring the SNR, SNDR, SFDR and ENOB etc.

1.3.2.1 Signal-to-Noise Ratio (SNR)

The SNR is the ratio of the ideal sinusoidal input signal power to the noise power. It is illustrated as [2]:

$$\begin{aligned} \mathrm{SNR} &= \frac{\mathrm{P_s}}{\mathrm{P_n}} = \frac{(\Delta \cdot 2^{N-1})^2/2}{\Delta^2/12} = 1.5 \cdot 2^{2N} \\ &= 6.02 \cdot N + 1.76 \ \text{(dB)} \end{aligned} \tag{1.5}$$

From the equation above, it is noted that SNR(dB) is increased by 6 dB for every additional bit [2].

1.3.2.2 Spurious Free Dynamic Range (SFDR)

The spurious free dynamic range (SFDR) is the ratio of the signal power to the strongest spurious signal within a certain frequency band. It is expressed in dBc as [2]:

$$\begin{aligned} \text{SFDR(dBc)} &= 10 \cdot \log\left(\frac{\text{Signal Power}}{\text{Largest Spurious Power}} \right) \\ &= 10 \cdot \log\left(\frac{A_{f_i}^2}{A_s^2} \right) \end{aligned} \tag{1.6}$$

where A_{fi} is the root-mean-square (rms) value of the fundamental input signal and A_s is the rms value of the largest spurious signal.

In effect, SFDR indicates the lowest-energy input signal that can be distinguished from spurious signals. Any signal below the SFDR cannot be reliably identified as a true signal instead of as a spurious one.

1.3.2.3 Total Harmonic Distortion (THD)

The non-linearities of the ADC generate spectral components at frequencies multiple of f_i (harmonics) [1]. The THD is the ratio between the power of all harmonics and the power of the input signals. It can be obtained as [1],

$$\begin{aligned} \text{THD(dB)} &= 10\log\left(\frac{\text{Total Harmonics Power}}{\text{Input Signal Power}}\right) \\ &= 10\log\left(\frac{\sum_{n=2}^{N_H+1} A_{nf_i}^2}{A_{f_i}^2}\right) \end{aligned} \tag{1.7}$$

where A_{n,f_i}^2 is the power of the spectral component at nf_i and N_H is the number of harmonics. A reasonable number for N_H is 9 [1].

1.3.2.4 Signal-to-Noise and Distortion Ratio (SNDR)

The SNDR is the ratio of the fundamental input signal power to the total noise and distortion power within a certain frequency band [2]. It is defined as:

$$\text{SNDR}=\frac{\text{Signal Power}}{\text{Noise and Distortion Power}} \tag{1.8}$$

SNDR is more useful in the real world, since it takes into consideration of the harmonic distortion power. SNDR can also be expressed by SNR and THD,

$$\mathrm{SNDR(dB)} = -10\log\left(10^{\frac{\mathrm{THD(dB)}}{10}} + 10^{-\frac{\mathrm{SNR(dB)}}{10}}\right) \tag{1.9}$$

This could be easily proved as below. First consider the SNR,

$$\begin{aligned} \mathrm{SNR(dB)} &= 10\log\frac{A_{f_i}^2}{A_{n1}^2} \\ A_{n1}^2 &= A_{f_i}^2 \cdot 10^{-\frac{\mathrm{SNR(dB)}}{10}} \end{aligned} \tag{1.10}$$

where A_{n1}^2 is the noise power. And according to previous equation, the total harmonics are $A_{n,H}^2 = A_{f_i}^2 \cdot 10^{\frac{\mathrm{THD(dB)}}{10}}$. Therefore, the SNDR is calculated as,

$$\begin{aligned} \mathrm{SNDR(dB)} &= 10\log\left(\frac{A_{f_i}^2}{A_{n1}^2 + A_{n,H}^2}\right) \\ &= 10\log\left(\frac{1}{10^{-\frac{\mathrm{SNR(dB)}}{10}} + 10^{\frac{\mathrm{THD(dB)}}{10}}}\right) \\ &= -10\log\left(10^{-\frac{\mathrm{SNR(dB)}}{10}} + 10^{\frac{\mathrm{THD(dB)}}{10}}\right) \end{aligned} \tag{1.11}$$

1.3.2.5 Effective Number of Bits (ENOB)

Another useful and straightforward parameter is the ENOB, which is related with SNDR. It is expressed as [2]:

$$\mathrm{ENOB} = \frac{\mathrm{SNDR(dB)} - 1.76}{6.02} \tag{1.12}$$

The ENOB indicates the effective number of bits which an ideal ADC could achieve according to its SNDR. As one may think, ENOB is definitely smaller than ideal bits.

1.3.2.6 Dynamic Range (DR)

The range from the full scale to the smallest detectable signal (SNDR=0) is called the dynamic range (DR) [2]. It is calculated as:

$$\begin{aligned} DR &= \frac{\text{Max Signal Power}}{\text{Min Signal Power (SNR=0dB)}} \\ &= \frac{\text{Max Signal Power}}{\text{Noise Power}} \end{aligned} \tag{1.13}$$

Normally, it can be expressed as the ratio of maximum signal power to the noise floor power.

CHAPTER 2

OVERVIEW OF HIGH SPEED ADC

2.1 Flash AD Converters

Flash converters are the fastest ADCs among other types of AD converters. It can achieve extremely high speed with low resolution. The typical structure of a flash ADC is shown in Figure 2.1.

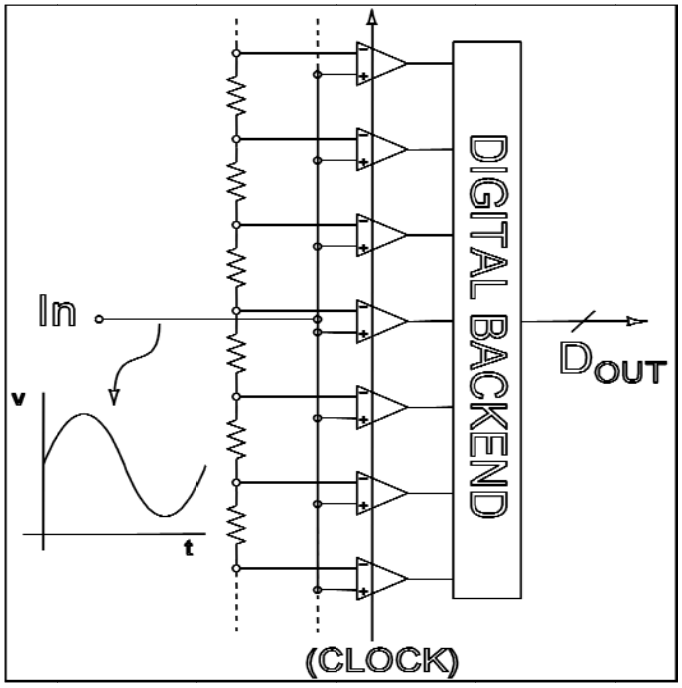


Figure 2.1 A Typical Structure of a Flash ADC.

An n-bit flash ADC consists of a resistor ladder composing 2^n equal resistors, 2^n-1 comparators and digital encoders. The resistor ladder is applied to produce different levels of references. Comparators are composed of a pre-amplifier used for sampling and amplifying the input signal, and a latch used for decision. Figure 2.2 shows the block diagram of a comparator. The encoder is of course applied to output the digital codes.

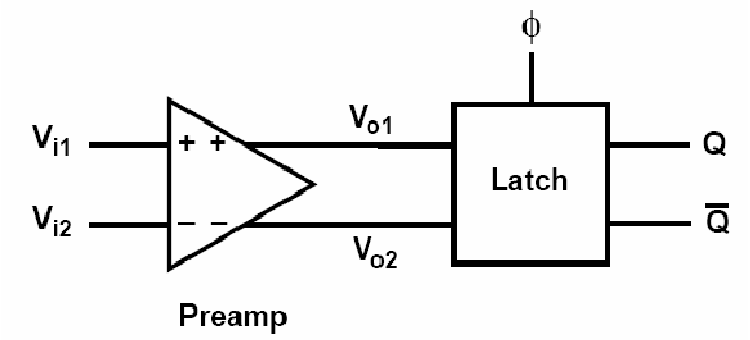


Figure 2.2 Block Diagram of a Comparator.

Each comparator samples the input signal and compares this signal with the reference voltage. Then the comparator generates a digital output "1" or "0" indicating whether the input signal is larger or smaller than the reference voltage assigned to that comparator. The digital outputs of the comparators are often referred to as thermometer codes, and these codes are encoded as 1-of-N codes. Then, 1-of-N codes are converted to Grey codes to suppress sparkle code and metastability, and finally Grey codes are converted to binary outputs. Figure 2.3 illustrates the encoding process.

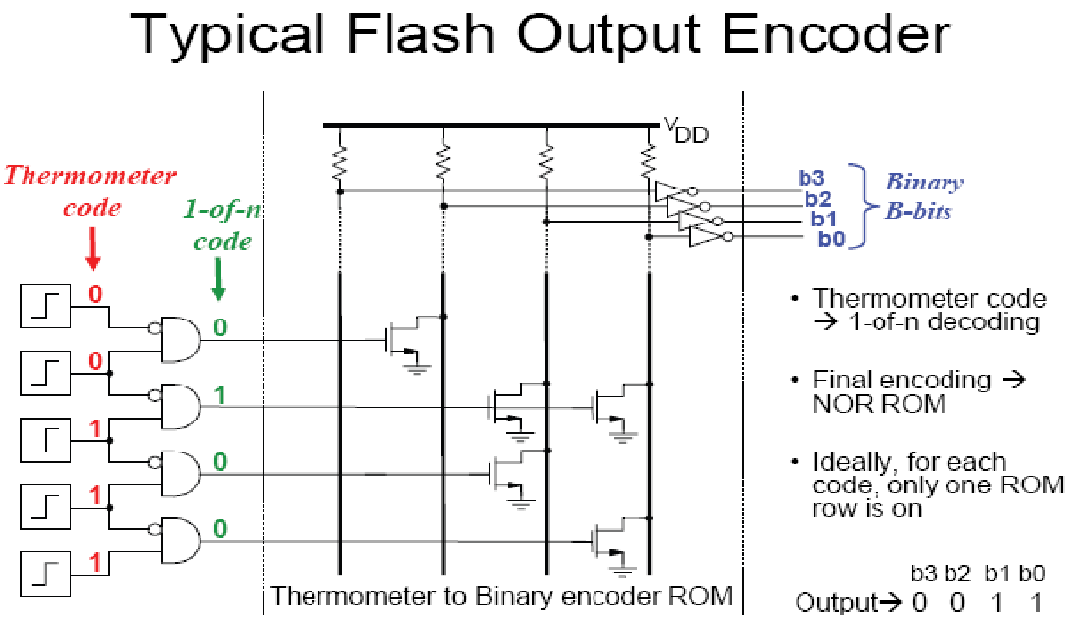


Figure 2.3 A Flash Output Encoder.

Despite of the high speed that a typical flash ADC can reach, it suffers a couple of drawbacks due to the large number of comparators and the lack of a front-end sample-and-hold circuit. Since the number of comparators and the resistors grows exponentially with the resolution, the flash ADC requires extremely large power dissipation and area for resolution above 5 bit. Furthermore, the large number of comparators gives rise to problems such as dc and ac deviation of the reference voltages generated by the ladder, large nonlinear input capacitance, and kickback noise at the analog input. On the other hand, the lack of a front-end sample-and-hold circuit makes the converter difficult to sample the input signal when the input signal has a very high speed. Thus, for modern flash ADC design, a sample-and-hold circuit is definitely required, although this will

result in other issue, like how to design a fast enough and accurate enough sample-and-hold circuit.

2.2 Two-Step Flash ADC

The exponential growth of power dissipation, area, and input capacitance of a flash convert makes it impractical for resolution above 5 bits. Therefore, trades-offs should be made between the resolution and the conversion rate. A two-step flash ADC is applied to trade speed for power dissipation and resolution.

A two-step flash ADC consists of a coarse flash ADC stage, a DAC, a subtractor and a fine flash ADC stage. Normally, a front-end sample-and-hold circuit, an inter-stage gain amplifier between the subtractor and the fine flash ADC are necessary. The block diagram in Figure 2.4 [3] illustrates the structure of a two step flash ADC.

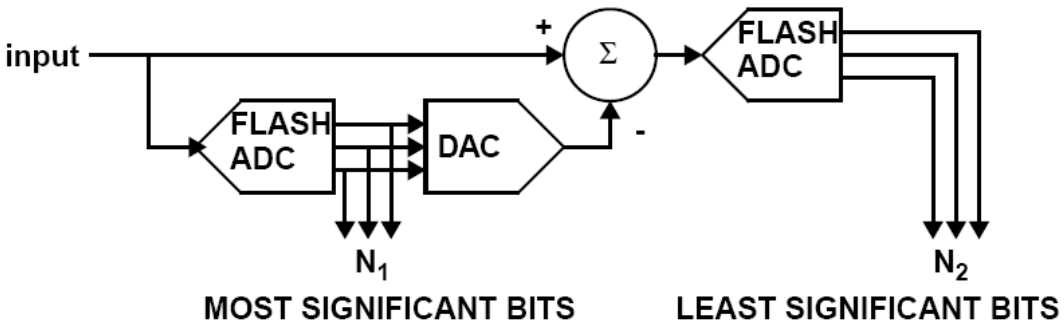


Figure 2.4 A Two-Step Flash ADC.

In this type of ADC, the conversion does not happen all at once as in the flash ADC. Here, the conversion takes two steps. During the first step, the most significant bits of the digital output are determined by the first stage flash ADC. Then a DAC converts this digital result back to an analog signal to be subtracted from the input signal. This residue is amplified by the inter-stage gain amplifier and then sent to the second stage flash ADC. The second stage flash determines the least significant bits of the digital output.

The key principle of a two-step flash ADC is to amplify the residue of the coarse ADC, and this will largely reduce the number of comparators. The number of comparators used in a two-step flash is only $2^{N1}+2^{N2}-2$ instead of $2^{N1+N2}-1$ comparators used in a straight flash ADC.

For example, a 4 bit two-step flash ADC is working as illustrated in Figure 2.5.

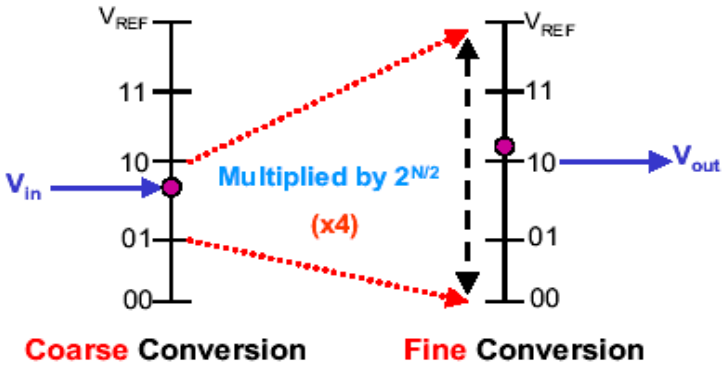


Figure 2.5 Principle of a Two-Step Flash ADC.

2.3 Folding and Interpolation Converters

2.3.1 Interpolation

The large input capacitance, high power dissipation, stringent timing requirement, and large area of full-flash architectures have limited the applications. A number of circuit techniques have been proposed to alleviate these problems while maintaining the one-step nature of the architecture [4]. Among these techniques, interpolation is quite preferred. Figure 2.6 [6] shows a simple resistive interpolated flash ADC, where the highlighted part is the interpolation block.

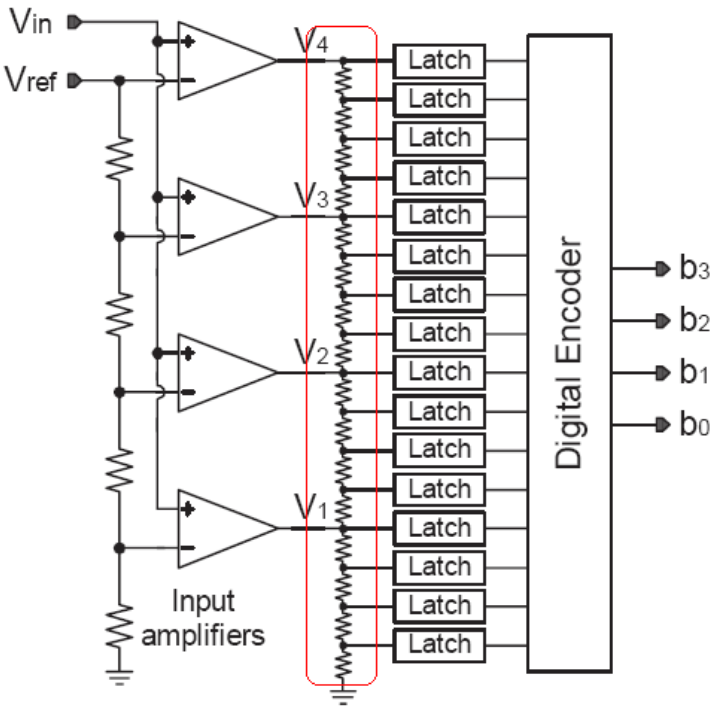


Figure 2.6 An Interpolated Flash ADC.

In order to reduce the number of the pre-amplifiers of the comparators at the input of the flash ADC, the difference between the analog input and each reference voltage can be quantized at the output of each pre-amplifier. This is illustrated by means of Figure 2.7 [4].

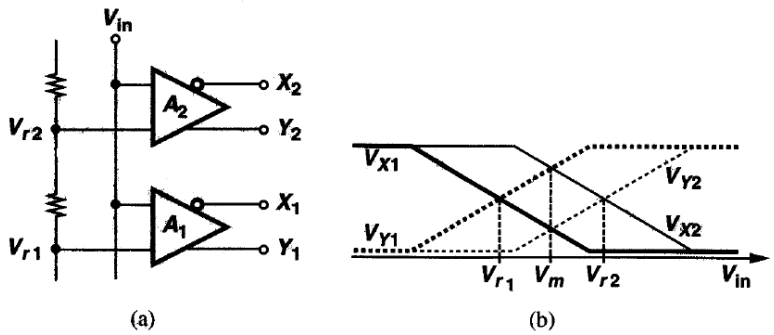


Figure 2.7 Principle of Interpolation Technique.

In Figure 2.7(a) we can see that pre-amplifier A_1 and A_2 compare the analog input with V_{r1} and V_{r2}, respectively. In Figure 2.7(b), the input/output characteristics of A_1 and A_2 are shown. Assuming zero offset for both pre-amplifiers, it is noted that $V_{x1}=V_{y1}$ if $V_{in}=V_{r1}$ and $V_{x2}=V_{y2}$ if $V_{in}=V_{r2}$. More importantly, $V_{x2}=V_{y1}$ or $V_{x1}=V_{y2}$ if $V_{in}=V_m=(V_{r1}+V_{r2})/2$; i.e., the polarity of the difference between V_{x2} and V_{y1} or V_{x1} and V_{y2} is the same as that of the difference between V_m and V_{in} [4]. If $V_{in}<V_m$, then $V_{y1}<V_{x2}$ or $V_{y2}<V_{x1}$; if $V_{in}>V_m$, then $V_{y1}>V_{x2}$ or $V_{y2}>V_{x1}$. Therefore, it is doable to compare V_{y1} with V_{x2} (or V_{y2} with V_{x1}) instead of comparing V_{in} with V_m (the middle value voltage).

The above observation indicates that the equivalent resolution of a flash stage can be increased by interpolating between the outputs of two adjacent pre-amplifiers. Figure

2.8 shows how an additional latch detects the polarity of the difference between single-ended outputs of two adjacent pre-amplifiers.

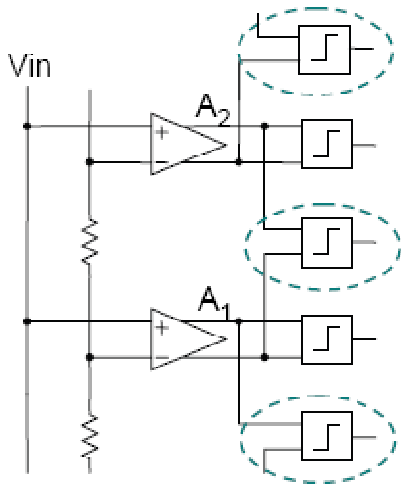


Figure 2.8 Interpolation Applied in a Flash ADC.

Note that in contrast with a simple flash stage, this approach halves the number of preamplifiers but maintains the same number of latches. From the figure above, the interpolation factor is 2. Generally the pre-amplifier burns more power than the latch. By this way the power dissipation is reduced. More important thing is that the input capacitance is reduced which will relieve the drivability of the input signals and the reference voltages are also largely reduced, where less front pre-amplifiers are used.

The concept of interpolation can be extended so as to produce more quantization levels between every two consecutive reference voltages in a flash converter, further reducing the number of input pre-amplifiers. For instance, Figure 2.9 [4] illustrates an interpolation factor of 4.

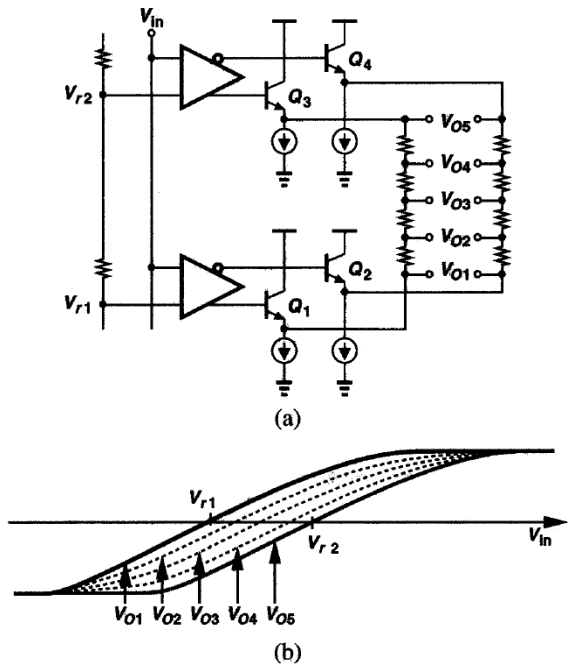


Figure 2.9 An Interpolation with a Factor of 4.

The outputs of two pre-amplifiers are interpolated using two uniform resistor strings. As illustrated in figure 2.9(b), since the input/output characteristics of the two pre-amplifiers are offset by V_{r2}-V_{r1}, as V_{in} goes from below V_{r1} to above V_{r2}, the differential output voltages $V_{o1},...V_{o5}$ cross zero at $V_{in}=V_{r1}+k(V_{r2}-V_{r1})/4$, for k=0, … , 4, respectively. If latches are used to detect the polarity of $V_{o1},...,V_{o5}$, this configuration provides an interpolation factor of 4. Thus, at least four times less reference voltages and front pre-amplifiers (input pre-amplifiers) could be saved depending on the number of interpolated stages. Note that the number of latches is still the same as in a full flash architecture.

A critical design issue that should be considered is the interpolation factor. The resistor strings and the input capacitance of the following latches introduce a time constant in the signal path, thereby reducing the bandwidth. This reduction is proportional to the square of the interpolation factor and hence becomes substantial if this factor exceeds around 4. Normally, the interpolation factor is ranging from 2 to 4.

2.3.2 Folding

Folding architectures evolve from full-flash and two-step flash topologies. As discussed before, a full-flash architecture provides a one-step operation but suffers from large hardware requirement, input capacitance, power dissipation and timing problems. On the other hand, a two-step architecture largely reduces the comparators but adds a front-end S/H circuit and analog postprocessing. Folding architectures, taking the advantage of both structures, apply less hardware while maintaining the on-step nature of the flash architecture. Figure 2.10 illustrates the block diagram of a folding ADC.

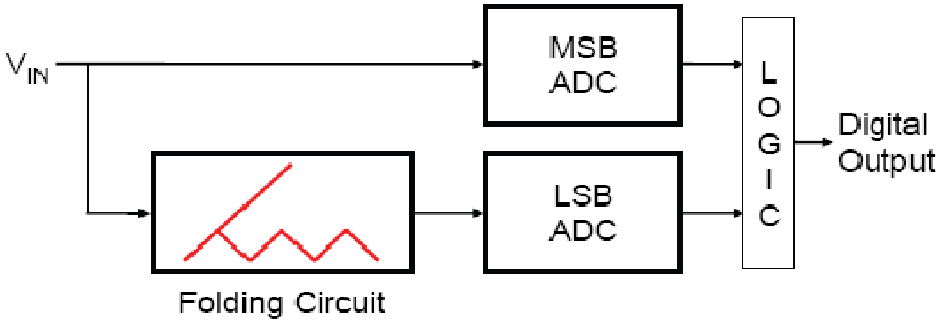


Figure 2.10 A Folding ADC.

The basic principle of folding is to generate a residue voltage through the folding circuit and subsequently digitize that residue to obtain the least significant bits (LSB). The MSB can be resolved using a coarse flash ADC operating in parallel with the folding circuit and hence samples the signal at approximately the same time that the residue is sampled. So the folding ADC is a one step ADC actually. A typical folding circuit is shown in Figure 2.11.

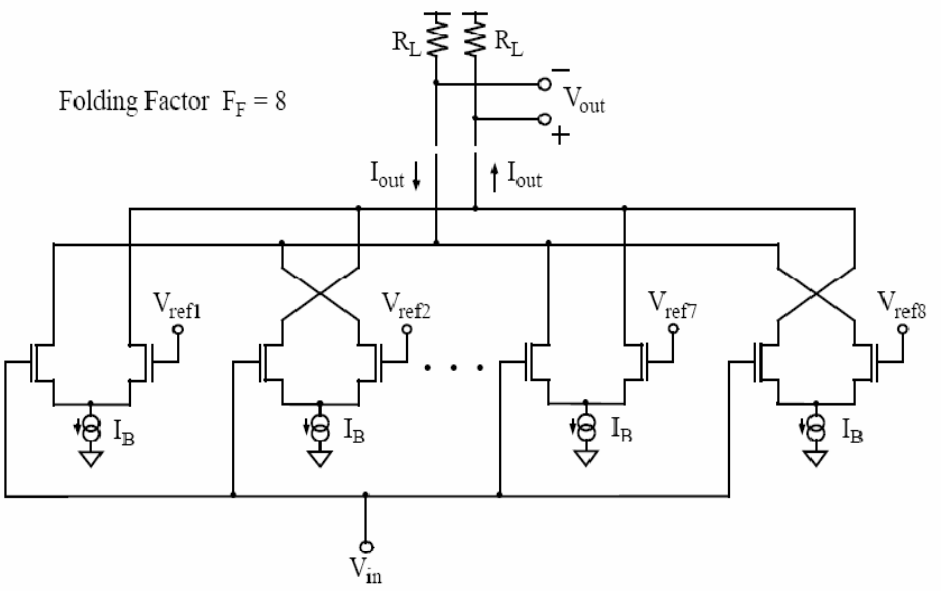


Figure 2.11 A Typical Folding Circuit.

The residue quantized by the fine flash ADC is generated by the folding circuit, which has the transfer function shown in Figure 2.11 [1].

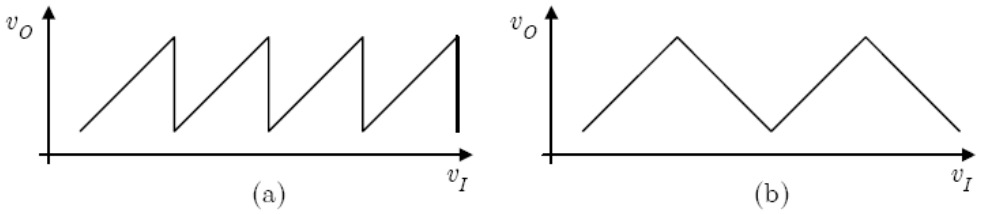


Figure 2.11 Transfer Function of Folding Circuits:
(a) Conceptual; (b) Low Speed in Practical.

However, in a practical high speed scenario, the actual transfer function of a folding signal is shown in Figure 2.12 [1]. The comparison of actual and real function of a folding circuit and its non-linearity errors are also illustrated in Figure 2.13 [8].

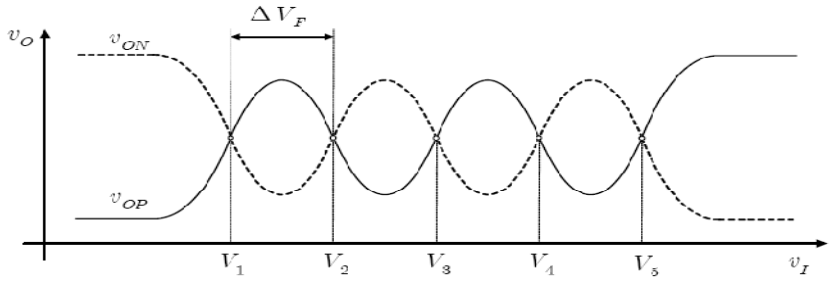


Figure 2.12 Transfer Function of Practical High Speed Folding Circuits.

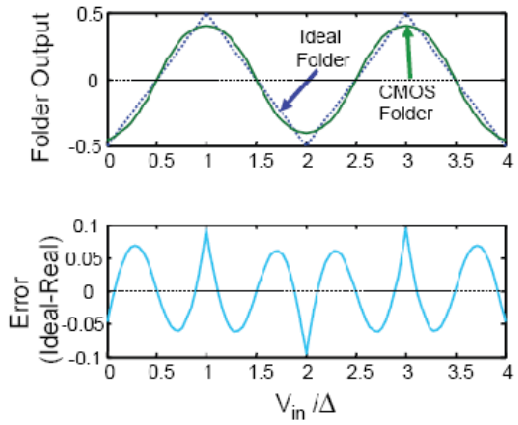


Figure 2.13 Outputs of Folding Circuit and Non-linearity Errors.

From the figure above, it is noted that the linearity in the wave crest is poor. Only the zero cross points are accurate. In order to overcome this drawback, parallel folding circuit, also called distributed folding circuit, is implemented in the real design, shown in Figure 2.14 and 2.15 [8, 11].

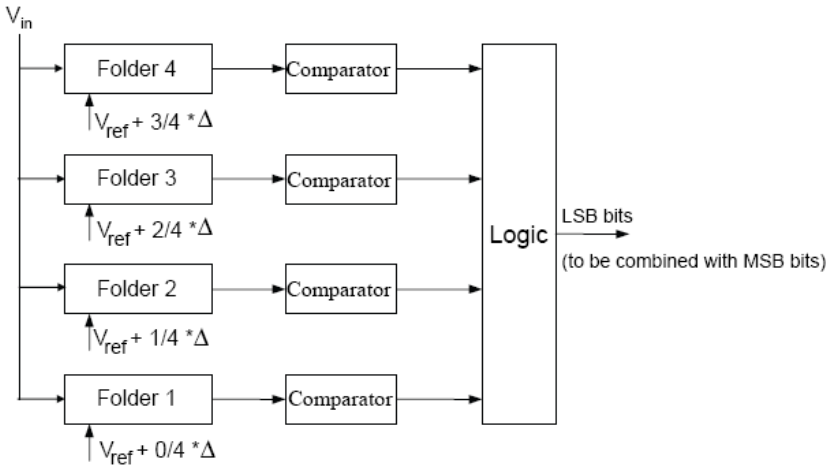


Figure 2.14 A Block Diagram of a Parallel Folding Circuit.

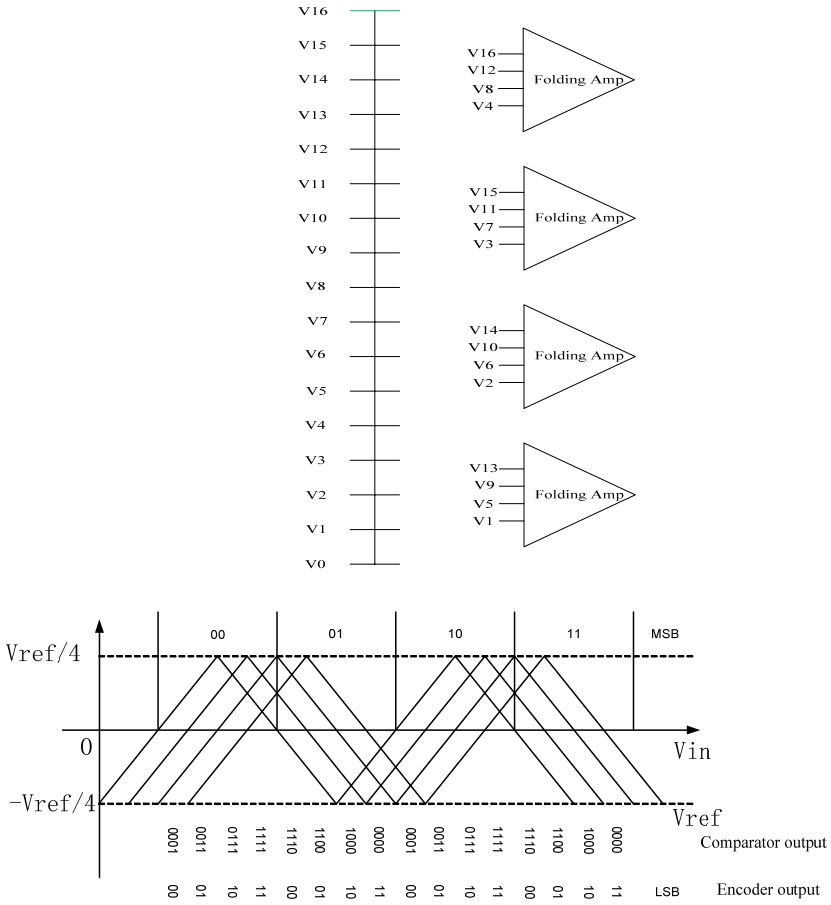


Figure 2.15 Parallel Folder Using only Zero-crossings.

It is manifested that more zero cross points are generated through the way that the input signals are feed into multi parallel folding amplifiers [11]. Only the zero cross points are used for the comparators to generate the output code. The codes come from the fine ADC is periodic circler codes which can be transfer to binary code while he output codes from the coarse ADC are just thermometer code.

If higher bit resolution or more zero-crossings are required, resistive interpolation technique can be applied. Figure 2.16 [11] illustrates a folding ADC combined with

resistive interpolation technique. As discussed before, interpolation factor is normally equal to or less than 4 and resistive interpolation will decrease the gain.

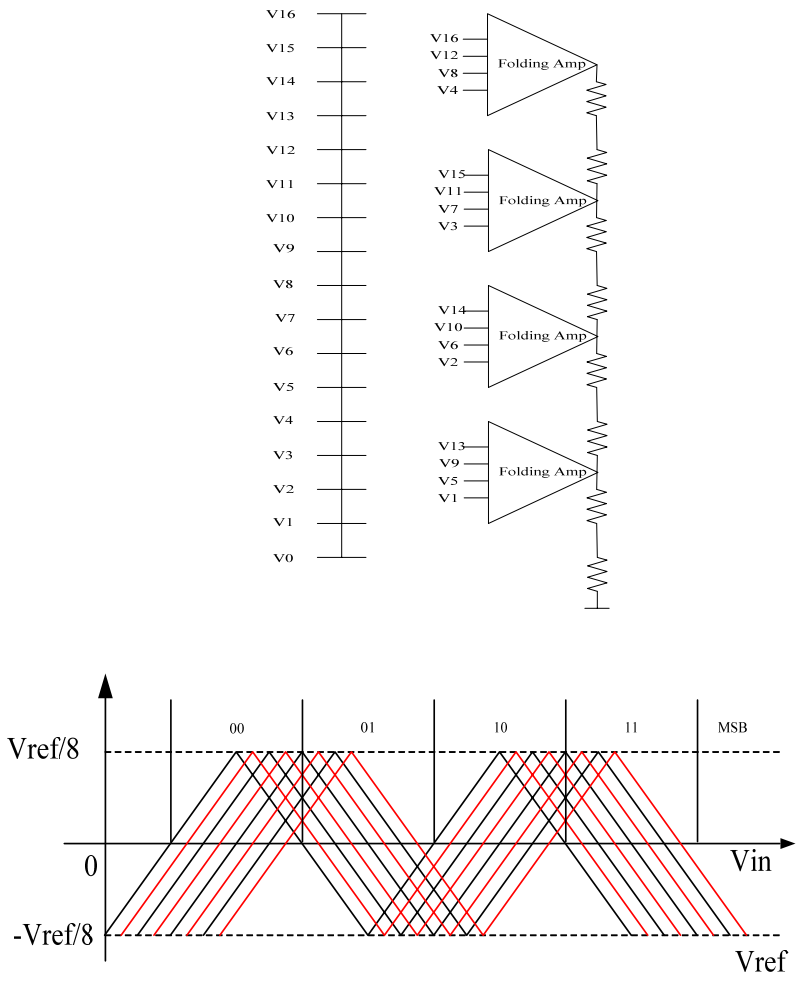


Figure 2.16 Folding Circuits Combined with Interpolation.

Previously, folding circuits obtained all the desired folding signals in a single step. Because of the interpolation factor is usually less than 4, more parallel folders are need to get enough zero cross point. Generally there is no enough headroom for so many parallel folders. In this case, the cascade folding structure can be used.

Previous folding structure indeed increases the bit resolution while still maintaining relatively high sampling speed. But there are also several limitations [1]:

1. The load capacitance rises with folding factor, due to the increasing number of differential pairs connected to the output, which limits the maximum sampling frequency;
2. Zero-crossing deviations increase and gain decreases due to the large folding factor and the not completely unbalanced differential pairs;
3. The common mode of the output voltage decreases if folding factor increases.

Therefore, the cascaded folding technique is introduced to mitigate those issues, allowing reaching a bit resolution as high as 13 bits, while maintaining a reasonably small number of differential pairs [1]. Figure 2.17 [1] shows an example of a cascaded folding circuit, where the 1st stage folding circuit has a folding factor of 5 and the 2nd stage folding circuit has a folding factor of 3. Thus the overall cascaded folding circuit has a folding factor of 15.

Cascaded folding technique makes it possible to generate larger folding factor with relative small interpolation factor because the folding factor can be distributed in different stages. More importantly, the interpolated resistor can be distributed in different stages. For example, a 2X interpolation in the pre-amplifier, a 2X interpolation in the 1st folding stage and a 2X interpolation in the 2nd folding stage will produce an 8X interpolation, where each of the folding factor is less than 4.

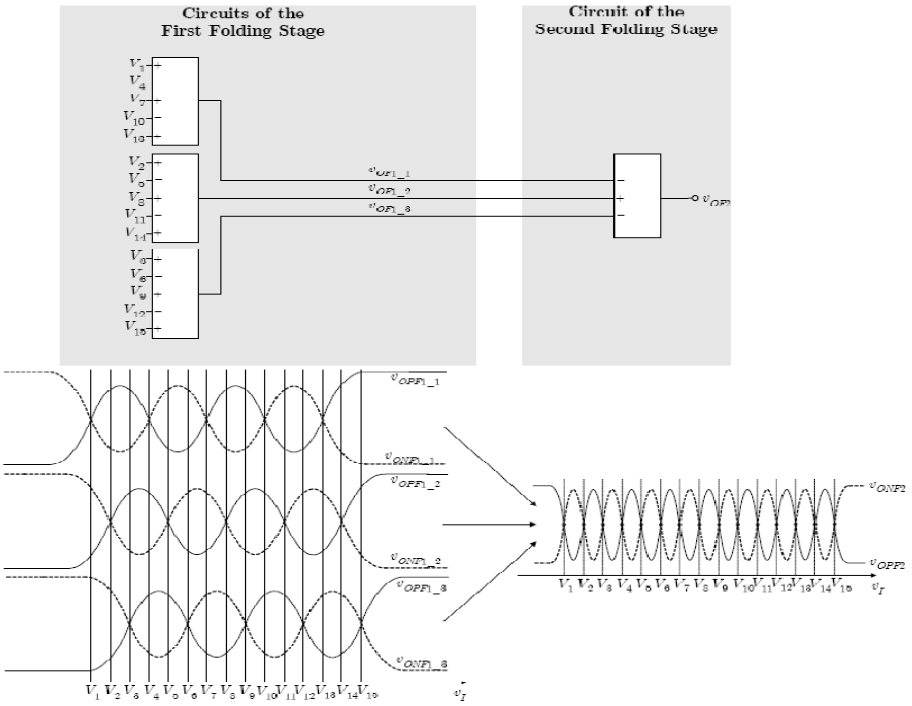


Figure 2.17 A Two-stage Cascaded Folding Circuit.

A more general N stage cascaded folding circuit is presented in Figure 2.18, where $F_F[k]$, $F_B[k]$ and $I_F[k]$ are folding factor, number of folding circuits and interpolation factor of k^{th} folding stage [1].

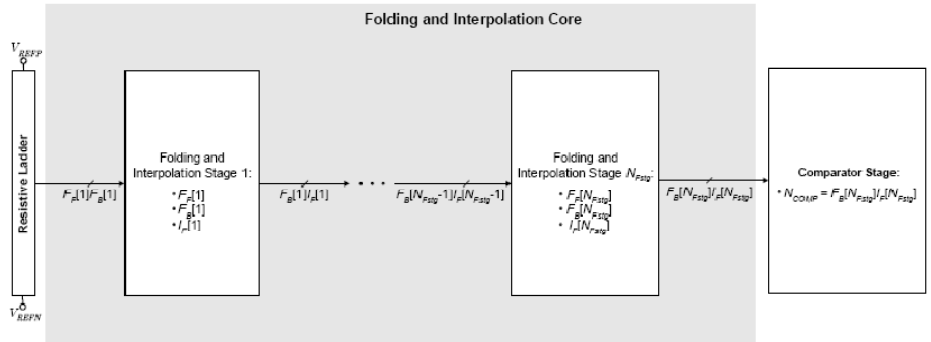


Figure 2.18 An N-stage Cascaded Folding Circuit.

2.4 Pipelined ADC

Pipelined ADCs are one of the most popular architectures for high resolution (over 10 bits up to 16 bits) and relative high speed (ranging from a few MSps to hundreds of MSps) applications. Pipelined ADCs have multiple cascades stages, compared to the two-step flash ADC which has just two stages. Each stage of the pipelined ADC consists of a sample-and-hold circuit, a sub-ADC, a DAC, a subtractor and an inter-stage gain amplifier. The block diagram of a pipelined ADC is illustrated in Figure 2.19.

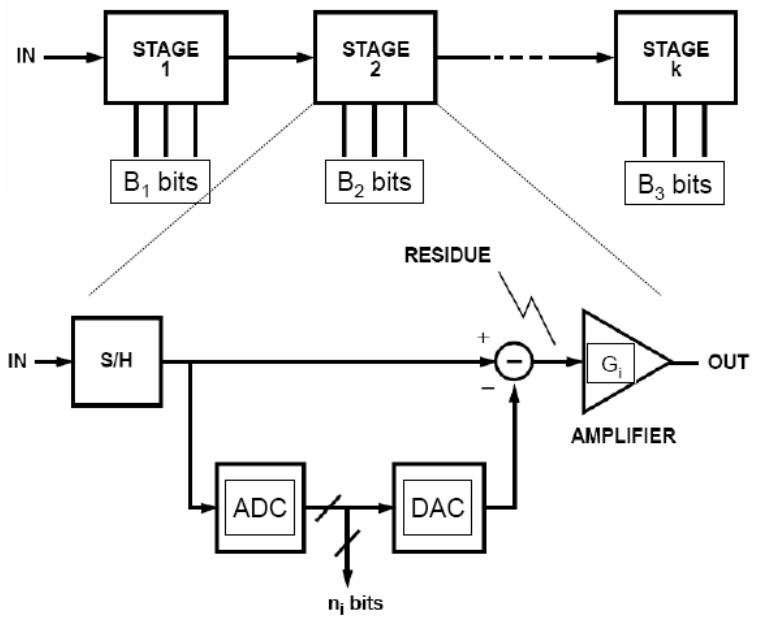


Figure 2.19 A Pipelined ADC.

The first stage sample-and-hold circuit samples the analog input signal, and then the sub-ADC converts the analog signal to digital outputs, which are called the most significant bit (MSB); meanwhile, the DAC converts the digital outputs back to an analog signal to be subtracted from the input signal, and this residue is amplified by the inter-

stage gain amplifier and sent to the following stage. The following stage samples the amplified residue and performs the same operation. The last stage contains only a sub-ADC and outputs the least significant bit (LSB). Therefore, all stages are working in this way. Note that all stages operate concurrently.

Only two phases of clock are needed in the pipelined ADC. During phase 1, all the odd stages samples the signals from previous stage, while all the even stages hold the circuits and convert the held signals from last phase; during phase 2, it operates the opposite way, all the odd stages converting signals and even stages sampling signals. This is shown in Figure 2.20 [8].

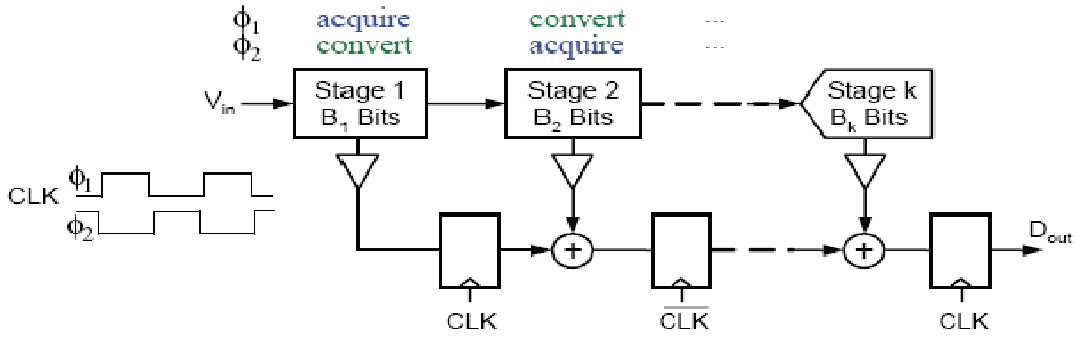


Figure 2.20 Working Process of a Pipelined ADC.

The concurrency of the pipelined ADC makes the maximum conversion rate almost independent of the number of stages because the first stage determines the conversion rate, but there is a delay time since the signal must work through all stages before the complete digital outputs are generated. This delay could be an issue if the

pipelined ADC is part of a feedback system [3]. In addition, the number of stages does have great impact on the noise performance, power dissipation, linearity and accuracy.

As a matter of fact, the stage partition of a pipelined ADC is considered as the most important and critical job. Stage partition may vary depending on different considerations, like noise, power consumption and area etc. A proper stage partition can reach a good performance and vice versa. Therefore, to design a pipelined ADC is a trade-off job among noise performance, conversion rate, power dissipation and etc.

2.5 Summary of ADCs

The previous contents illustrate several different kinds of ADCs used in the industry. Each type of ADCs has its own advantages and disadvantages. It is very important to choose the proper ADC in practical designs. For examples, Ultra-Wide Band (UWB) systems are applying flash ADCs in order to reach high speed while image sensor systems are using pipelined ADCs because of the high resolution requirement.

Table 2.1 shows the specification ranges of different ADCs.

	Flash	**Folding and Interpolation**	**Pipeline**	**Sigma-Delta**
Resolution	Less than 6 bits	6-8 bits	8-14 bits	Over 14 bits

Sampling Rate	Up to 5-6GSps in CMOS	Up to 1GSps	Up to 200MSps	Up to 10MSps
Power	Hundreds mW to several W	Hundreds mW	Hundreds mW	Tens of mW to hundreds mW

Table 2.1 Summary of Different Types of ADCs.

CHAPTER 3

4-BIT 2.3 GSPS ADC DESIGN

3.1 High Speed ADC Requirements

For a MOSFET, the highest intrinsic frequency that can be obtained is called the characteristic frequency. This characteristic frequency is normally considered how fast a MOSFET can work ideally and is expressed as,

$$f_T = \frac{g_m}{2\pi C_{gs}} = \frac{3}{4\pi}\frac{\mu}{L^2}\left(V_{gs} - V_T\right) \text{ if a MOSFET in saturation region.}$$
$$\text{or } f_T = \frac{v_{sat}}{2\pi L} \text{ if a MOSFET in velocity saturation.} \tag{3.1}$$

From the equations above, it is obvious that reducing the channel length or increasing the overdrive voltage will increase the characteristic frequency f_T, and the channel length is having larger impact than the overdrive voltage. Another interesting thing is that for velocity saturation case, characteristic frequency f_T has nothing to do with the overdrive voltage and is only inverse proportional to L instead of L^2.

Those two phenomenon illustrates that technology scaling down is good for increasing the ADC speed since the channel length decreasing, although the overdrive voltage inevitably keeps going down. This is correct whatever a MOSFET is in saturation or velocity saturation. On the other hand, it is not comprehensive to think technology scaling down is absolutely good for high speed ADCs. Due to the scaling, the power

supply is also decreasing, which means the input and output ranges are decreasing. If the input/output ranges are smaller, as offsets, mismatches and noises in short channel technologies behaving more severely, it is harder to overcome the undesired factors and hence reduces the overall performances, such as SNDR, ENOB, INL, DNL and etc.

Therefore, it is very much trade-off work to design a high-speed ADC in short channel technology. Higher speed may result in relatively poorer performance. However, even though there are some defects, technology scaling down is definitely the trend for high speed ADC designs and the only way known so far to increase the speed. More and more extreme high speed ADCs are under 90nm, 65nm or even 45nm CMOS technologies with a sampling rate reaching up to 10GSps.

3.2 Proposed Architecture

As introduced in Chapter 2 before, the most used and proper architectures for high speed ADC (over GSps) are flash and interpolation/folding. This proposed low-resolution high-speed flash ADC (4 bit 2.3 GSamples/s) applies capacitive interpolation and offset cancellation, where offset cancellation is implemented by capacitive averaging technique since it has no edge effect. Besides, the averaging capacitors also function as interpolation devices. No S/H circuit is used, instead capacitors are applied and working with switched-capacitor circuits.

The specification is shown in Table 3.1.

	Parameter	Typical value
DC Accuracy	DNL	+/- 0.5LSB
	INL	+/- 0.8LSB
	No missing code	
Analog Input	Voltage Range	0.6Vpp
Digital Output	Output voltage	1.2V
	Coding	Binary
Power Supply	Analog part	1.2V
	Digital part	1.2V
Dynamic Performance	SNDR	22dB
	SNR	22dB
	SFDR	27dB
Conversion Rate	2.3GSps(90mW)	
S/H Circuit		No S/H Circuit

Table 3.1 Specification of a 4-bit 2.3-GSps Flash ADC.

3.2.1 Differential Switched-Capacitor Pre-amplifier

Switched-Capacitor circuits are very often used in modern CMOS designs. Accurate S/H circuits, operational amplifiers with accurate amplification, differentiators, integrators and etc. all apply switched-capacitor circuits.

In this flash ADC, a specific switched-capacitor pre-amplifier is designed. There are many advantages of applying it, which will be discussed in this section and the following sections. For high speed flash ADC, it is not possible to apply a separate S/H circuit in the front since the S/H circuit can't operate under high speed and will definitely reduce the overall speed. So normally, a simple capacitor is used as an S/H circuit in order to sample and hold the input signals. Combining the capacitor with pre-amplifier, a switched-capacitor pre-amplifier is constructed. It is not only for amplifying the signals (pre-amplifier) but also for sampling and holding input signals (S/H circuit). Other advantages, such as offset cancellation, averaging and capacitive interpolation, can also be provided by switched-capacitor pre-amplifiers.

Shown below is a differential switched-capacitor pre-amplifier, which is applied in the design.

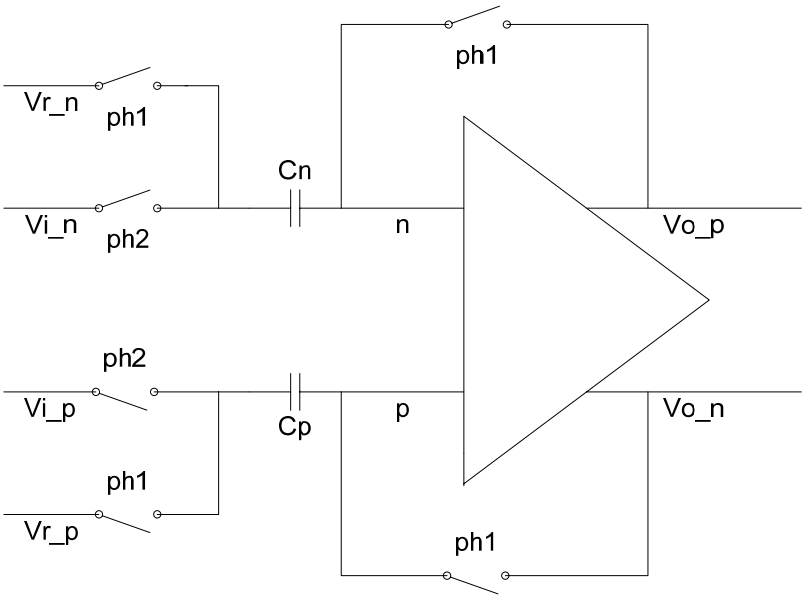


Figure 3.1 A Differential Switched-Capacitor Pre-amplifier.

where C_p and C_n (C_p=C_n) are sampling capacitors and *ph1* and *ph2* are two non-overlap control signals. The switched-capacitor pre-amplifier is working under two phases. First, take a look at the working mode in phase 1 as illustrated in Figure 3.2.

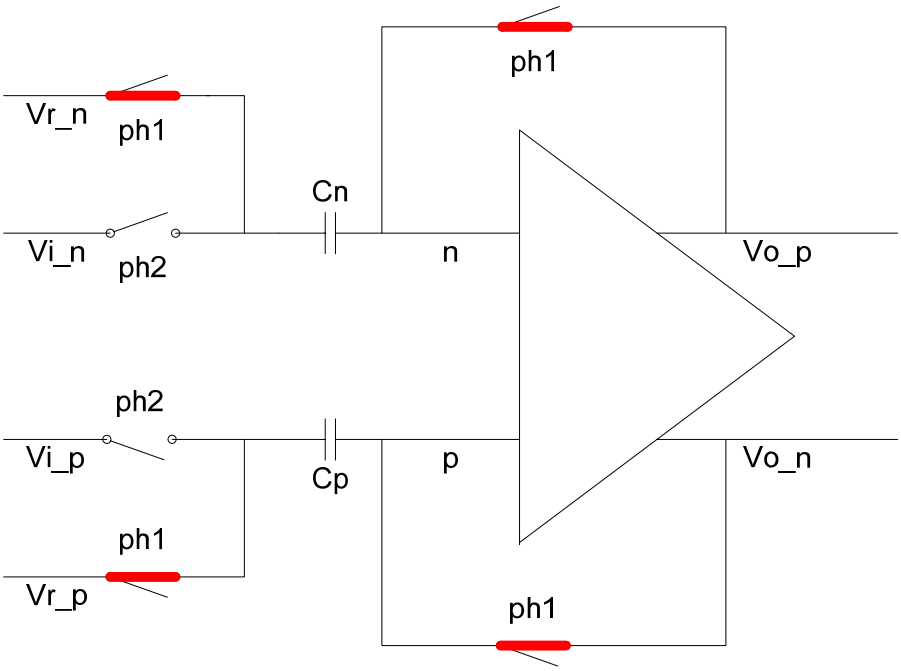


Figure 3.2 A Switched-Capacitor Pre-amplifier Working in Phase 1.

During *ph1*, four switches (highlighted in red) are closed and the pre-amplifier forms a closed-loop circuit with its input and output connected together. It is easily obtained that,

$$\begin{aligned} &\left(V_p - V_n\right)\cdot A_c = V_{o\_p} - V_{o\_n} \text{ and } V_{o\_p} = V_n,\ V_{o\_n} = V_p \\ &\text{Therefore } V_p = V_n = V \end{aligned} \tag{3.2}$$

Also consider the charge stored on both capacitors,

$$\begin{aligned} Q_n &= \left(V_{r\_n} - V_n\right)\cdot C_n = \left(V_{r\_n} - V\right)\cdot C \\ Q_p &= \left(V_{r\_p} - V_p\right)\cdot C_p = \left(V_{r\_p} - V\right)\cdot C \end{aligned} \tag{3.3}$$

This equation indicates that the differential reference voltages are applied and stored on both capacitors.

During *ph2*, the switched-capacitor pre-amplifier will take the input signals and amplify the differences between inputs and references. The working mode is shown below.

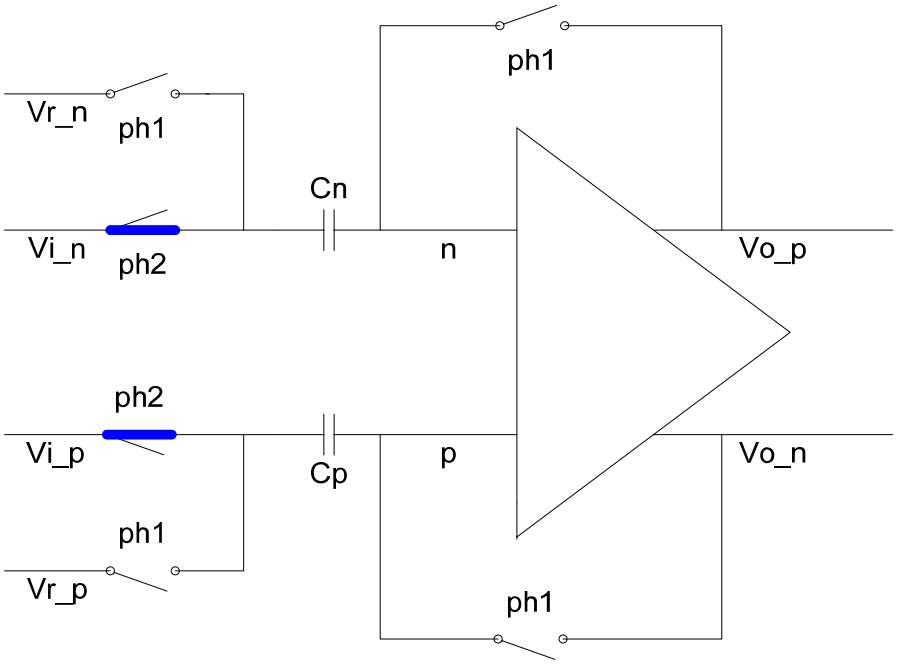


Figure 3.3 A Switched-Capacitor Pre-amplifier Working in Phase 2.

During *ph2*, *ph1* switches open and *ph2* switches closed, the differential input voltages are applied to both capacitors. Since the right plates of both capacitors are floated and no current allowed flowing, the charge on each capacitor will not change. So the following equation is obtained,

$$
\begin{aligned}
&Q_n^{'} = \left(V_{i\_n} - V_n^{'}\right)\cdot C = Q_n = \left(V_{r\_n} - V_n\right)\cdot C \\
&\Rightarrow V_{i\_n} - V_n^{'} = V_{r\_n} - V_n; \\
&Q_p^{'} = \left(V_{i\_p} - V_p^{'}\right)\cdot C = Q_p = \left(V_{r\_p} - V_p\right)\cdot C \\
&\Rightarrow V_{i\_p} - V_p^{'} = V_{r\_p} - V_p. \\
&\text{Thus, } V_p^{'} - V_n^{'} = \left(V_{i\_p} - V_{i\_n}\right) - \left(V_{r\_p} - V_{r\_n}\right) = \left(V_{i\_p} - V_{r\_p}\right) - \left(V_{i\_n} - V_{r\_n}\right)
\end{aligned}
\tag{3.4}
$$

Therefore, the outputs of the pre-amplifier can be expressed as,

$$
\begin{aligned}
V_{o\_p}^{'} - V_{o\_n}^{'} &= A\cdot\left(V_p^{'} - V_n^{'}\right) \\
&= A\cdot\left[\left(V_{i\_p} - V_{i\_n}\right) - \left(V_{r\_p} - V_{r\_n}\right)\right]
\end{aligned}
\tag{3.5}
$$

which clearly indicates that the differences between the input and reference signals are amplified. Because this pre-amplifier amplifies signals in open-loop, the working speed can be very high for a low gain.

The previous analysis shows that a switched-capacitor pre-amplifier can be applied to high speed flash ADCs acting as both an S/H circuit and a gain-stage amplifier.

3.2.2 Capacitive Interpolation

Interpolation technique, as introduced in Chapter 2, is mostly often used in high speed flash ADC design, because it reduces the input impedance and the number of references. There are typically two kinds of interpolation techniques, resistive interpolation and capacitive interpolation. Resistive interpolation uses resistors as the loads for pre-amplifier to generate zero-crossing points (details in Chapter 2). The interpolation factor of resistive interpolation could be any integer value theoretically

depending on how many segments of resistors are interpolated. Normally the factor ranges from 2 to 5. On the other hand, capacitive interpolation uses capacitors as the loads instead, and has a fixed interpolation factor of 2. A very typical structure of an interpolated ADC core with an interpolation factor of 2 is shown in Figure 3.4.

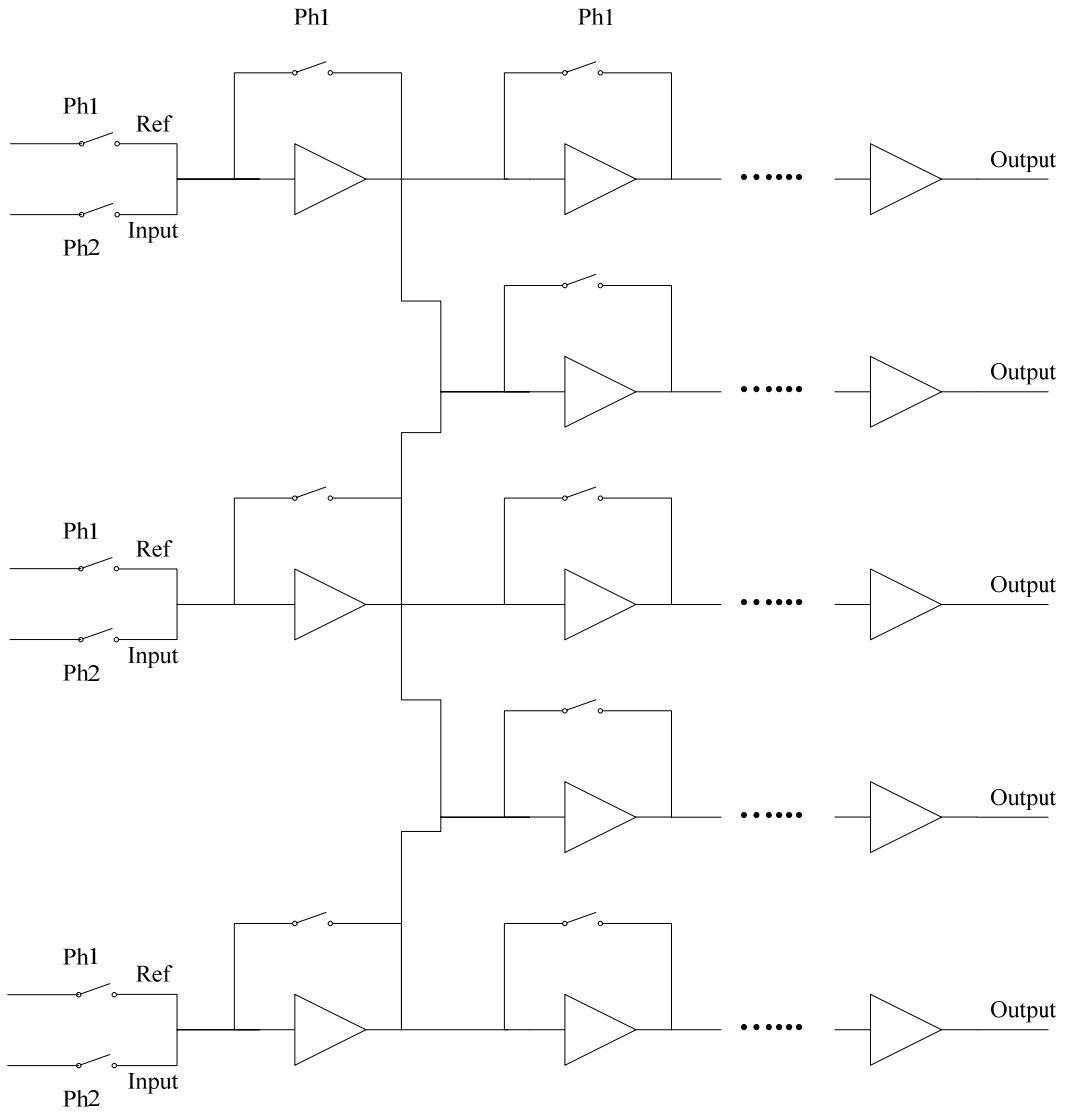


Figure 3.4 A Typical Interpolation Technique with Interpolation Factor of 2.

The two interpolation techniques have their own advantages and disadvantages. Compared with resistive interpolation, capacitive interpolation may apply several cascaded stages to reach the same required resolution bits because of its low interpolation factors, and will finally result in decreasing the overall ADC speed (analyzed in Chapter 4). However, capacitive interpolation is relatively easier to sample and hold the signals if switched-capacitor circuit is applied. Besides, the resistive interpolation could lower the pre-amplifier's gain and cause accuracy issues. Therefore, choosing the proper interpolation technique is very important and should depend on the practical specification.

In this thesis, capacitive interpolation is preferred because it uses a purely reactive averaging network between the outputs of the adjacent amplifiers [8]. It requires neither over range comparators which consumes a lot of power, nor any static averaging termination. Besides, external sample-and-hold circuit is not necessary since the interpolation capacitors used as distributed front-end S/H circuit at each stage can sample the signals [8].

The principle of capacitive interpolation is illustrated in Figure 3.5. All the capacitors including the front end and interpolated ones are acting as distributed S/H circuits. All the interpolated capacitors are acting as two purposes, one for interpolation and the other for averaging. The interpolated capacitors have capacitances only half of the non-interpolated ones. Because as we know before the voltage in the interpolated node is the middle voltage of two adjacent output voltages of the pre-amplifiers, in order

to charge the interpolated capacitor to the required voltage value within the same time period as others do, half capacitance of the interpolated capacitor is applied.

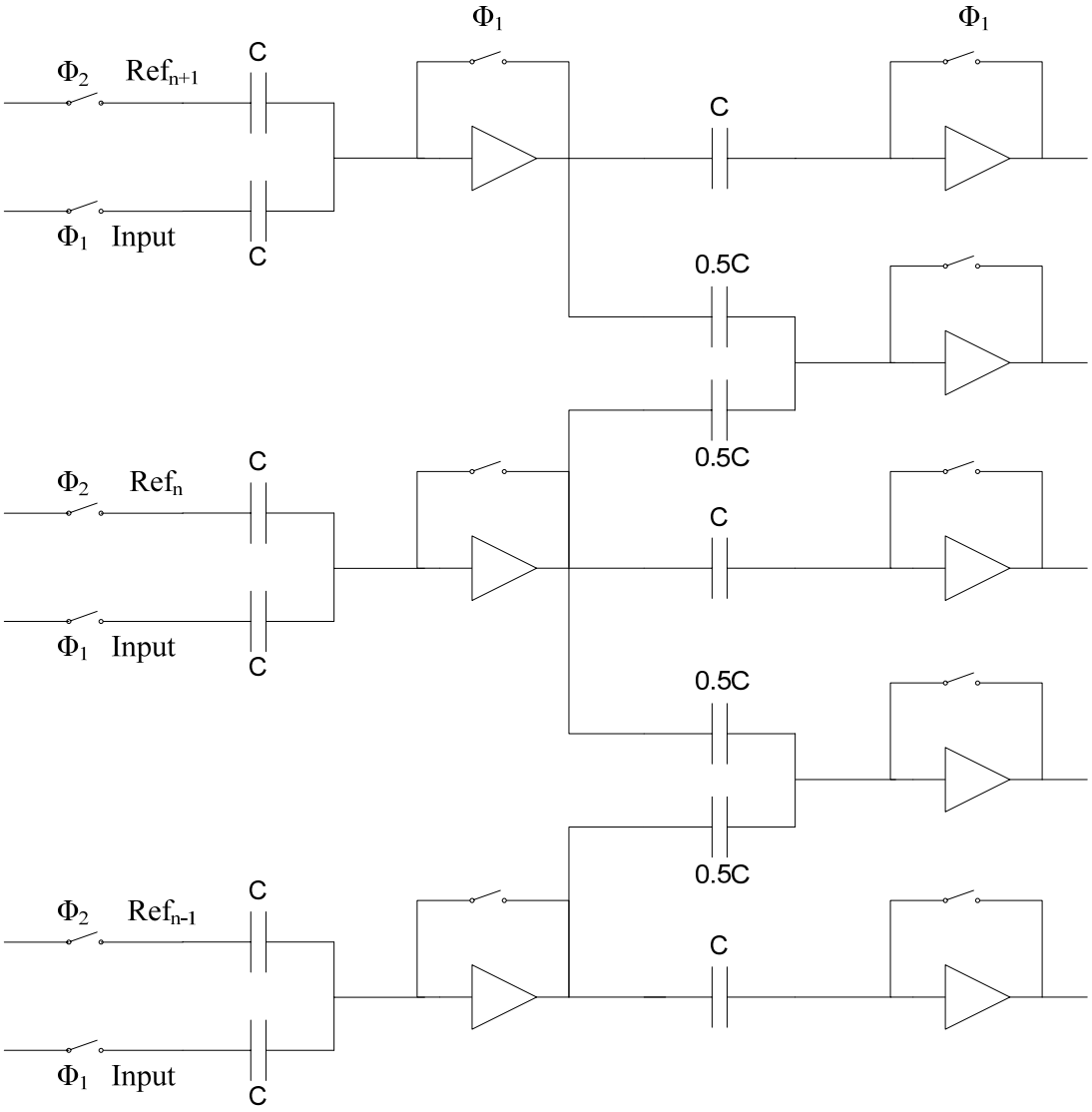


Figure 3.5 Capacitive Interpolation Techniques.

During phase Φ_1 the input signal is tracked and the amplifier works in the reset mode. At the end of the phase Φ_1 the input signal is sampled at the input capacitors.

During phase Φ_2 the reference voltage is tracked and then the voltage difference between the input voltage and the reference voltage is coupled to the input node of the amplifiers. At this period the amplifier works in the amplification mode so that the voltage difference is amplified.

From this figure it is noted that the capacitors interpolate the reference voltage. The middle amplifier compares the voltage difference between the input and the half value of the reference. So another zero crossing point can be generated. By the capacitance interpolation the static power is zero. That is another benefit [9].

One drawback of the capacitive interpolation is the capacitive divider and the input capacitance of the amplifier during the amplification phase [8]. The overall gain of each stage is given by the intrinsic gain of the amplifier ad the capacitive interpolation factor. In order to minimize the total input capacitance of the converter, the sampling capacitors will be chosen as small as possible.

3.2.3 Averaging Technique

For flash ADCs, as discussed before, if interpolation technique is applied, zero-crossing points will be generated. Ideally, the zero-crossings are equally placed as shown below [10].

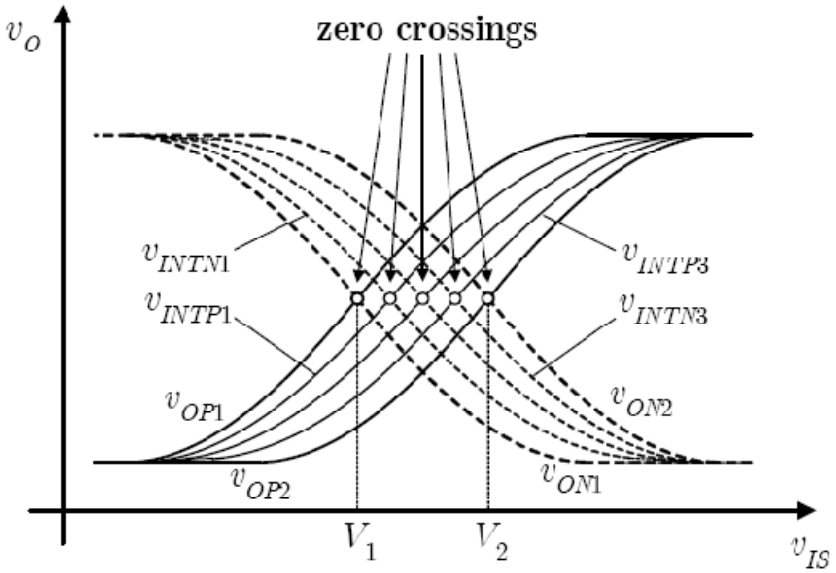


Figure 3.6 Zero-crossings Ideally Placed by Interpolation.

However, the non-linear transfer function of the pre-amplifiers will cause the deviation of the zero-crossings, which are away from their ideal locations, as illustrated below.

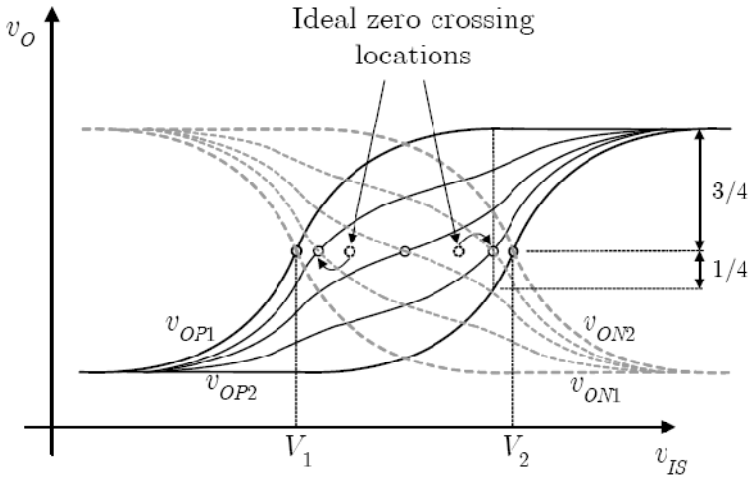


Figure 3.7 Zero-crossings Deviation Due to Non-linearity.

It is easy to understand that the deviations (offsets) are actually resulted from the mismatches of the pre-amplifiers and the loadings (resistors or active components). Moreover, the location of different zero crossings are uncorrelated random variables [10]. Therefore, utilization of averaging technique will reduce the offsets of the pre-amplifiers. The averaging network produces output voltages, which are determined by a weighted sum of the outputs of several differential pairs [10]. There are different kinds of averaging techniques, such as resistive averaging, capacitive averaging and etc. A general averaging network is shown below [10].

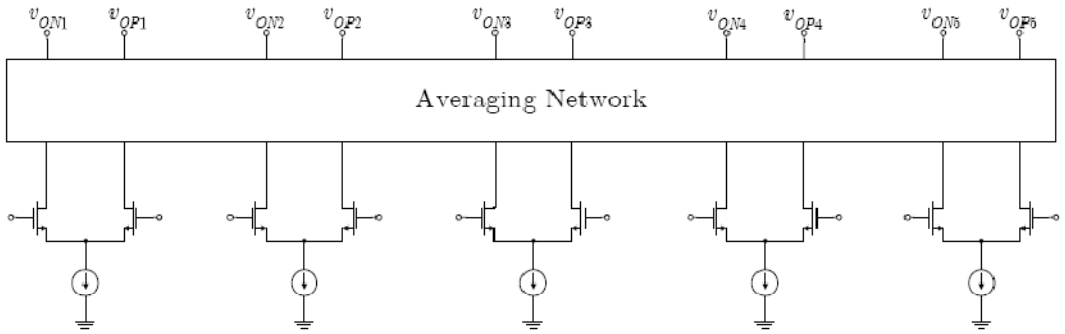


Figure 3.8 A General Averaging Technique Network.

Fortunately, interpolation and averaging techniques can both be implemented in one way, either resistively or capacitively. In this thesis, because capacitive interpolation is applied, capacitive averaging technique is used accordingly.

A capacitive averaging network is illustrated below.

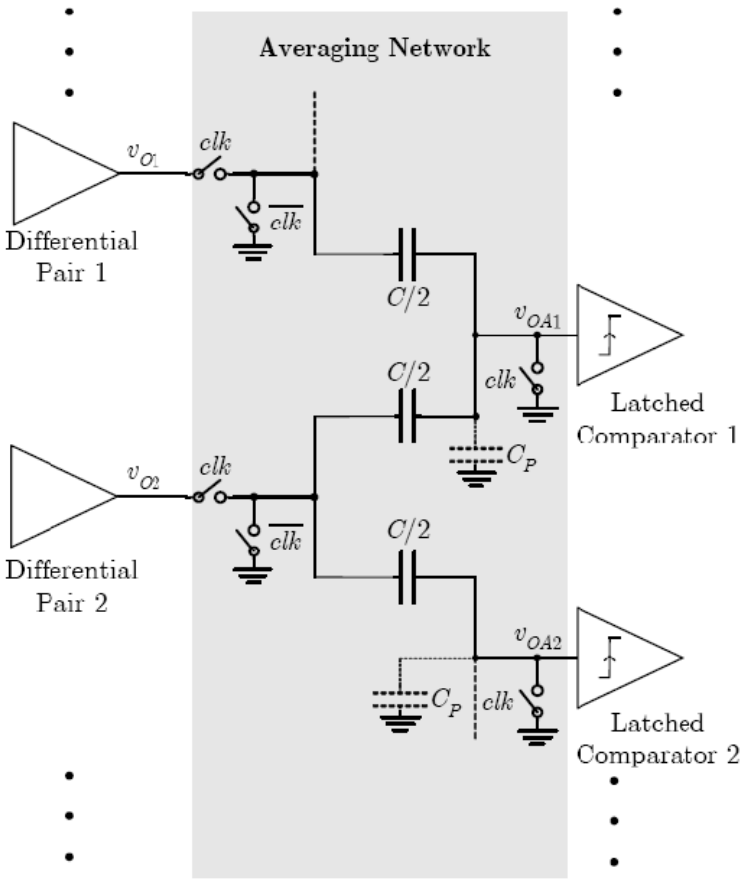


Figure 3.9 A Capacitive Averaging Network.

During *clk* = high, *clk* switches on, the output voltages of the pre-amplifiers are charged (sampled) to capacitors *C*/2. For node A1, two capacitors being connected together, so the total charge of the two is $Q = v_{O1} \cdot \frac{C}{2} + v_{O2} \cdot \frac{C}{2}$. During *clk* = low, *clk_b* switches on, a charge redistribution will occur, the total charge is then $Q' = v_{OA1} \cdot C + v_{OA1} \cdot C_P$. Since the charge doesn't change, so the input voltage of the latched comparator is obtained as

$$v_{OA1} = -\frac{C}{C + C_P} \cdot \frac{v_{O1} + v_{O2}}{2} \tag{3.6}$$

where C_P is the parasitic capacitance at node A1.

If the two pre-amplifiers have offset voltages V_{OS1} and V_{OS2} respectively, the offset in the input of latched comparator 1 is easily calculated as,

$$v_{OSA1} = -\frac{v_{OS1} + v_{OS2}}{2} \tag{3.7}$$

As illustrated above, the input offset is the average of V_{OS1} and V_{OS2}. Assuming the standard deviations of V_{OS1} and V_{OS2} are σ_{OS1} and σ_{OS2} respectively, and $\sigma_{OS1} = \sigma_{OS2} = \sigma_{OS}$, the following is obtained,

$$\sigma_{OSA1} = \frac{\sigma_{OS}}{\sqrt{2}} \tag{3.8}$$

Therefore, it is clearly shown that the standard deviation of the offset voltage is decreased because of averaging technique.

3.2.4 Offset Cancellation

In this ADC design, the offset cancellation technique is applied. First, let's take a look at the input offset storage (IOS). Figure 3.10 depicts a configuration employing this technique.

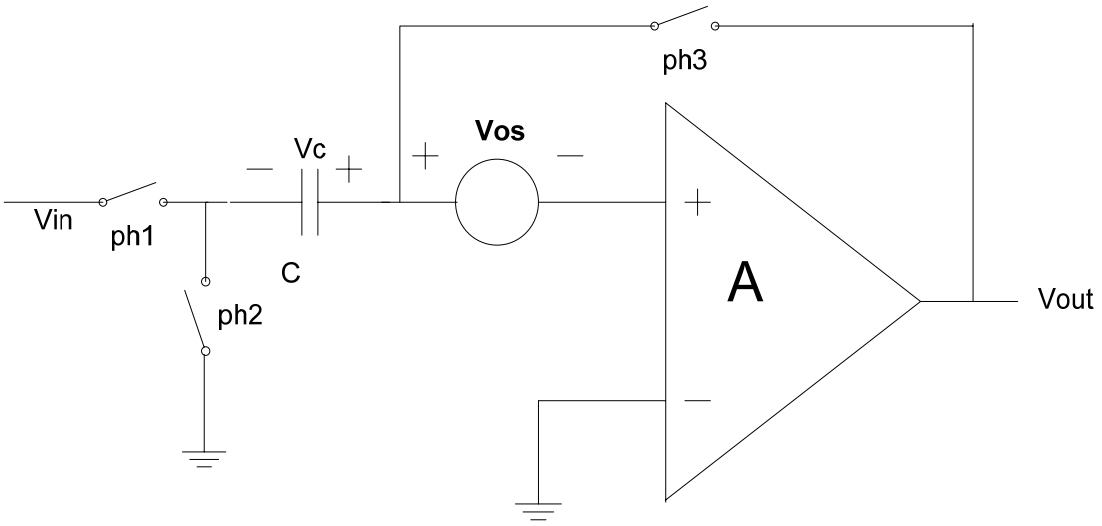


Figure 3.10 An Input Offset Storage Technique.

During Φ_1, ph_2 and ph_3 are closed, a unity gain feedback loop is established and the offset is stored in the capacitor. During Φ_2, ph_1 is closed and ph_2 and ph_3 are open, the input signal is added and amplified. In the offset cancellation mode [4],

$$\begin{aligned} \left(V_{out} - V_{OSA}\right)\left(-A_0\right) &= V_{out}, \\ V_{out} &= \frac{A_0}{1 + A_0} V_{OSA} \end{aligned} \tag{3.9}$$

Thus, the amplifier output offset is nearly V_{OSA}, in contrast to the no offset cancellation, output being $A_0 V_{OSA}$.

An improved way to cancel the offset is the output offset storage (OOS) shown in Figure 3.11.

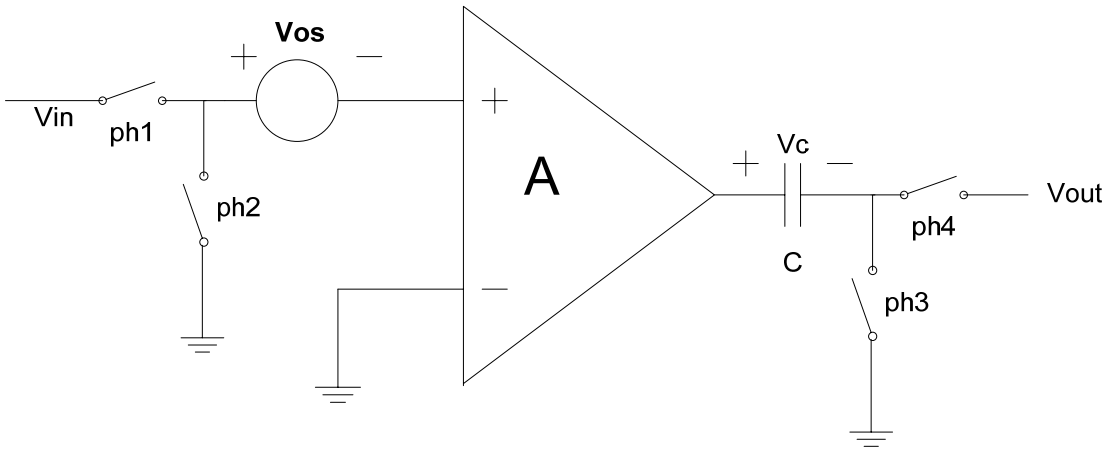


Figure 3.11 An Output Offset Storage Technique.

During Φ_1, ph_2 and ph_3 are closed, and the offset is amplified and stored. During Φ_2, ph_1 and ph_4 are closed, ph_2 and ph_3 are open, the output voltage is calculated as,

$$\begin{aligned} V_{out} &= A \cdot \left(V_{in} + V_{OS} \right) - V_C \\ &= A \cdot \left(V_{in} + V_{OS} \right) - A \cdot V_{OS} \\ &= A \cdot V_{in} \end{aligned} \tag{3.10}$$

Thus the output offset is $V_{OS(tot)} = \frac{\Delta q}{A_0 C} + \frac{V_{OSL}}{A_0}$.

In this thesis, since capacitive interpolation is applied, which requires cascaded pre-amplifier networks, the multi-stage offset storage technique is used. Figure 3.12 shows the structure in differential form.

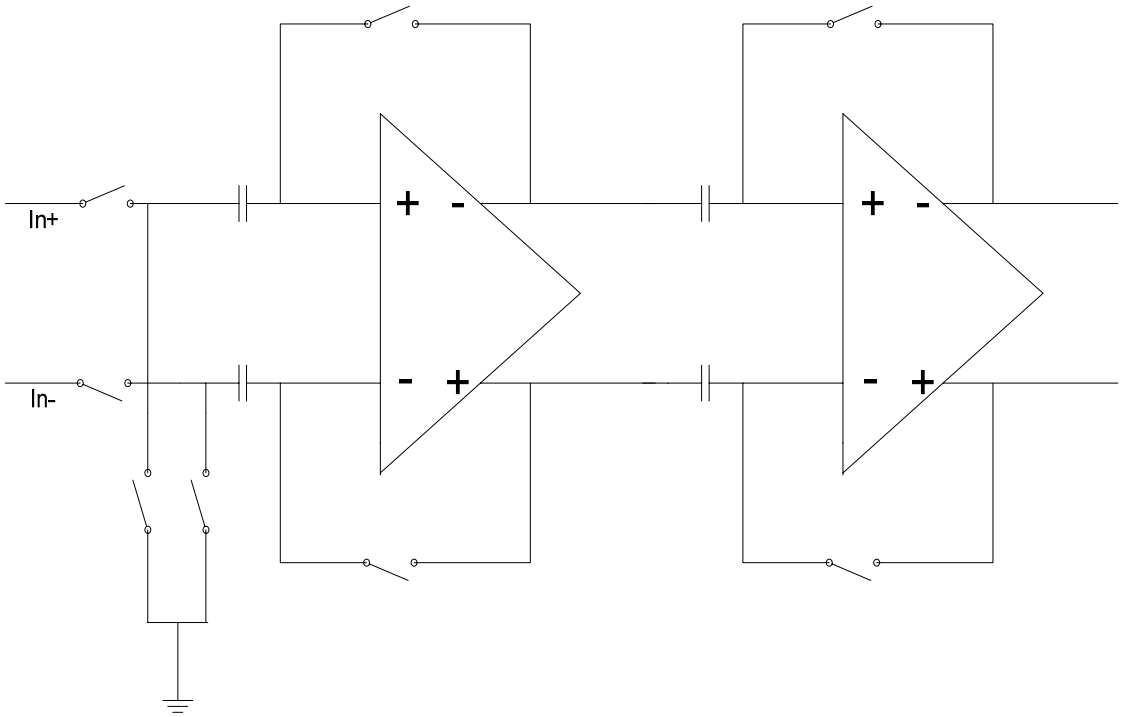


Figure 3.12 A Multi-stage Offset Storage Technique.

Do a quick analysis below. During Φ_1, the switches highlighted in Figure 3.13 are closed, and references and offsets voltages are both stored. Thus, it is obtained that $V_{C+} - V_{C-} = \left(V_{\mathrm{Re}f+} - V_{\mathrm{Re}f-}\right) - V_{offset}$.

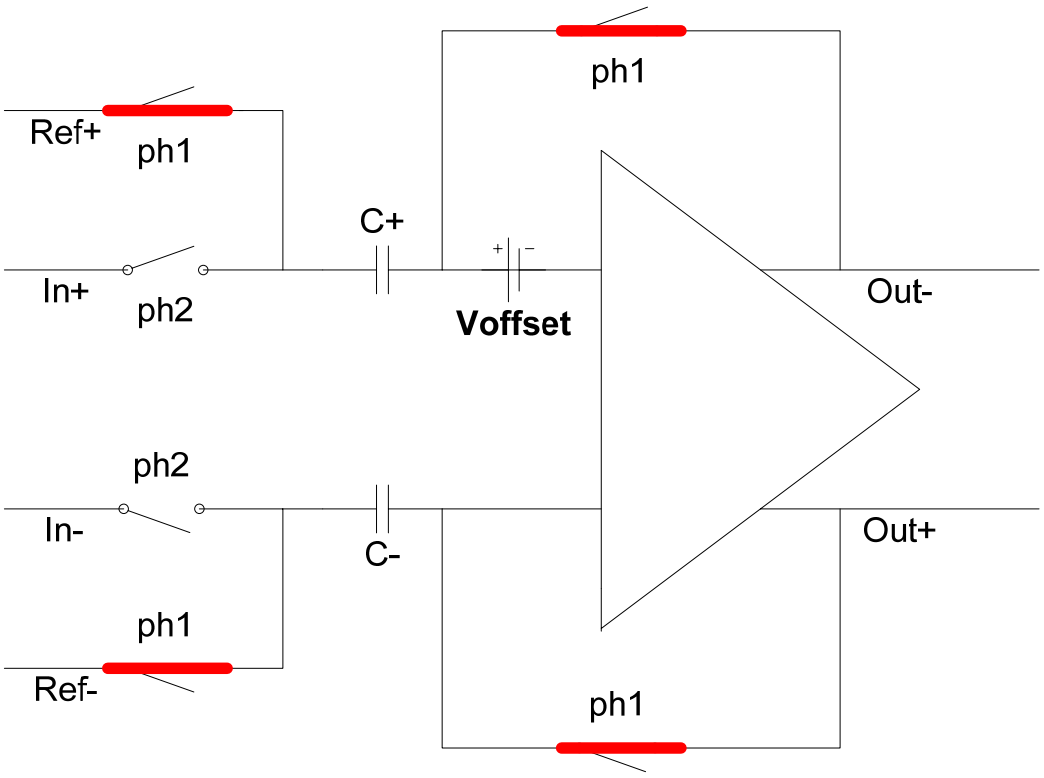


Figure 3.13 Offset Cancellation During Φ_1.

During Φ_2, the witches highlighted in Figure 3.14 are closed while other switches are open. so $V_o = A \cdot \left[\left(V_{in+} - V_{in-} \right) - \left(V_{C+} - V_{C-} \right) - V_{offset} \right]$. Substituting for V_{C+}-V_{C-}, it is calculated that $V_o = A \cdot \left[\left(V_{in+} - V_{in-} \right) - \left(V_{\mathrm{Re}f+} - V_{\mathrm{Re}f-} \right) \right]$. Therefore, the offset is cancelled and difference between input and reference is established.

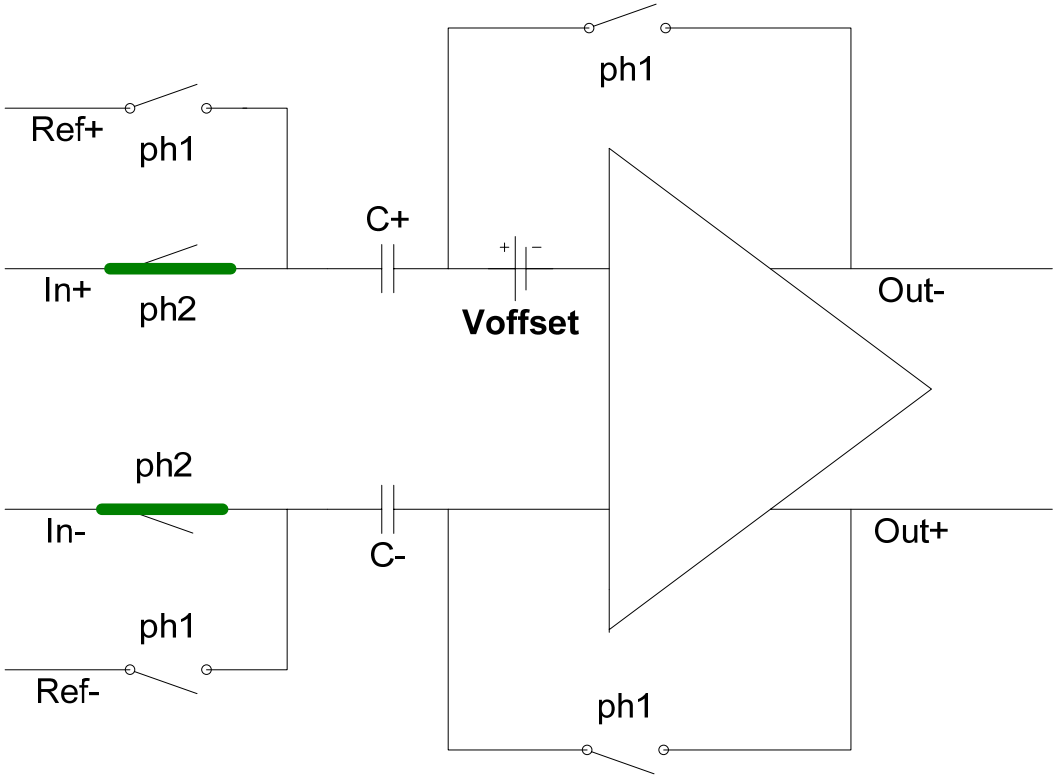


Figure 3.14 Offset Cancellation During Φ_2.

Multi-stage offset storage technique is quite often used in capacitive interpolation flash ADCs.

3.2.5 Comparators

A comparator is the last step of signal comparison, which determines the output a high voltage (VDD) or a low voltage (GND) before sending to the digital back-end, an encoder. For high speed ADCs, latched comparators are mostly applied because of the very fast comparison speed. There are mainly three types of latched comparators [10],

static latched comparator, Class-AB latched comparator and dynamic latched comparator. One common of all these latched comparators is that they are all positive feedback circuits, which thereafter result in very fast comparison. In this thesis, the dynamic latched comparator is used because it has the fastest comparison speed.

First, take a look at the static latched comparator, which was shown in Figure 3.15 [10].

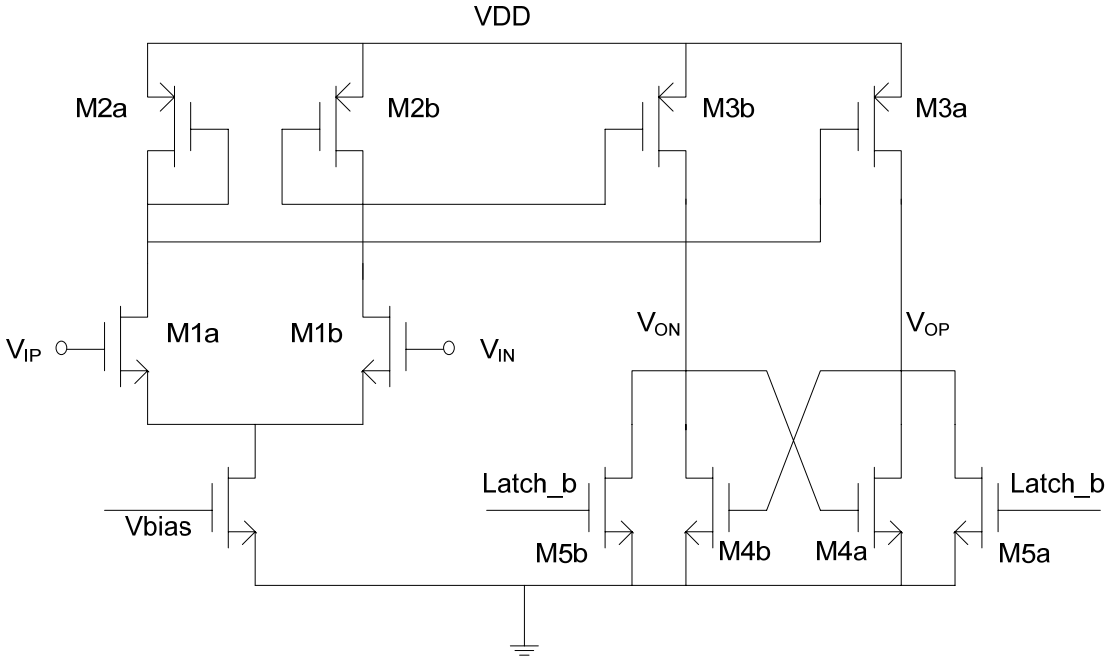


Figure 3.15 The Structure of a Static Latched Comparator.

During reset, *Latch_b* = high, M_{5a} and M_{5b} are on and push the outputs to ground. $M_{1a/b}$ and $M_{2a/b}$ are acting as a pre-amplifier, which mirrors the current through $M_{3a/b}$ [10]. During regeneration, *Latch_b* = low, $M_{5a/b}$ are off and the current starts flowing into $M_{4a/b}$. M_{4a} and M_{4b} are cross-coupled, forming a positive feedback loop. Depending on the input

voltages, either M_{4a} or M_{4b} will turn on first and initiates the regeneration process. If $V_{IN} > V_{IP}$, the current flowing through M_{3b} is larger, and M_{4a} will turn on first, thus finally push down V_{OP} to GND and V_{ON} to VDD.

The static latched comparator presents poor power efficiency because the pre-amplifier is always working and consuming static power. And the comparison speed is limited since there are two dominant poles, one at the output node of the pre-amplifier and the other at the output node of the regenerator [10]. However, the kickback noise is very small [10].

Another latched comparator is called Class-AB latched comparator, shown below in Figure 3.16. As the name illustrated, the regeneration is done by two cross-coupled Class-AB inverters.

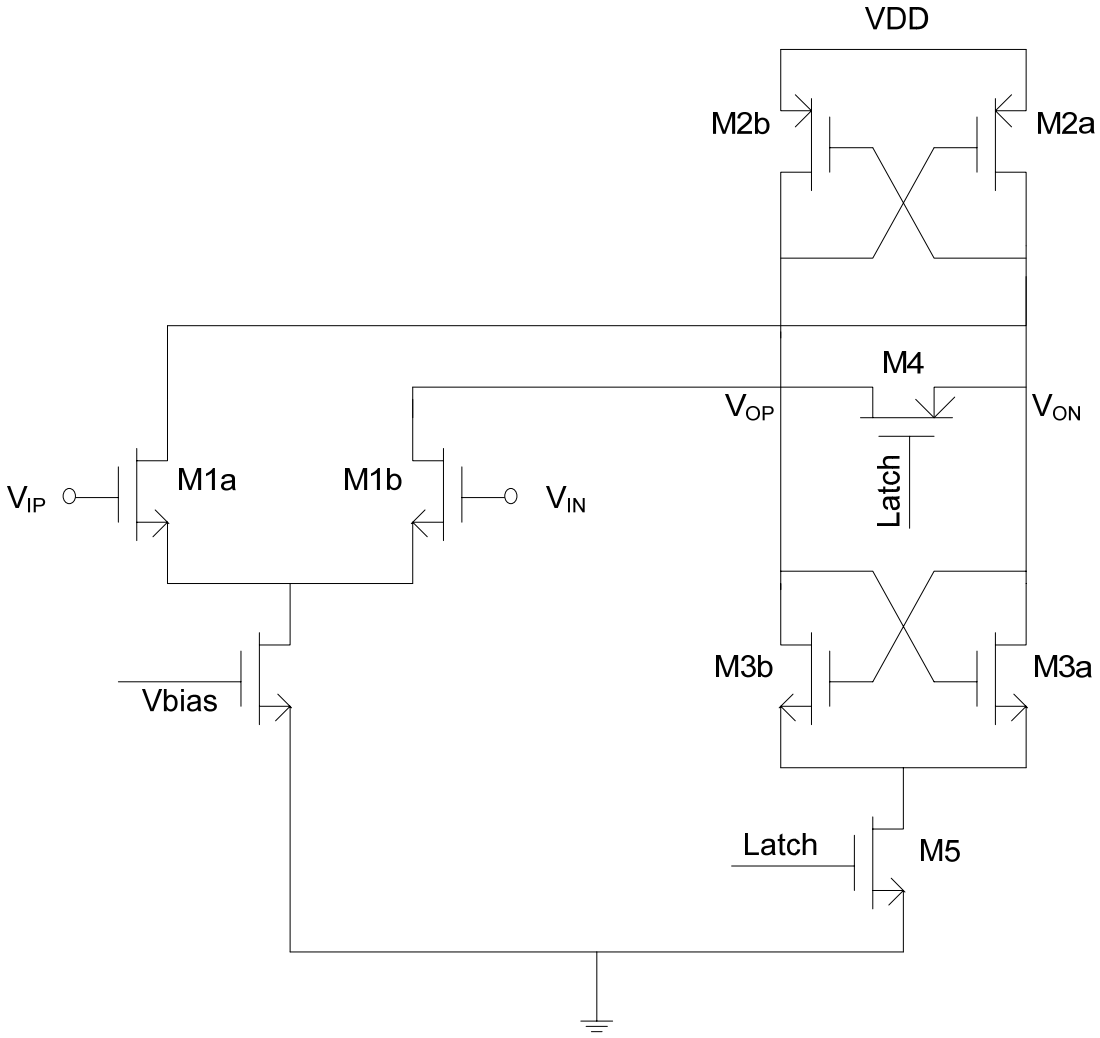


Figure 3.16 A Structure of Class-AB Latched Comparator.

During reset phase, *latch* = low, M_5 is cut-off and no current flows through $M_{3a/b}$ [10]. M_4, the reset switch, turns on, besides M_4 along with $M_{2a/b}$ together forms the loads to the pre-amplifier so that signals can be amplified. During regeneration phase, *latch* = high, M_4 turns off. M_{2a}/M_{3a} and M_{2b}/M_{3b} form two cross-coupled inverters, which are

positive feedback circuits. Therefore, the output signals will be either pushed up to VDD or pushed down to GND momentarily.

As clearly illustrated in the Class-AB latched comparator, there is only one pole which is located at the output nodes, so the circuit has a faster speed than static one (because of larger bandwidth). On the other hand, because of rail-to-rail connections at output nodes, the output voltages will be capacitively coupled to inputs [10] and of course disturbing the inputs, which causes larger kickback noise.

The third type as well as the fastest and the most power efficient is the dynamic latched comparator, which is applied in this thesis. The structure is shown in Figure 3.17.

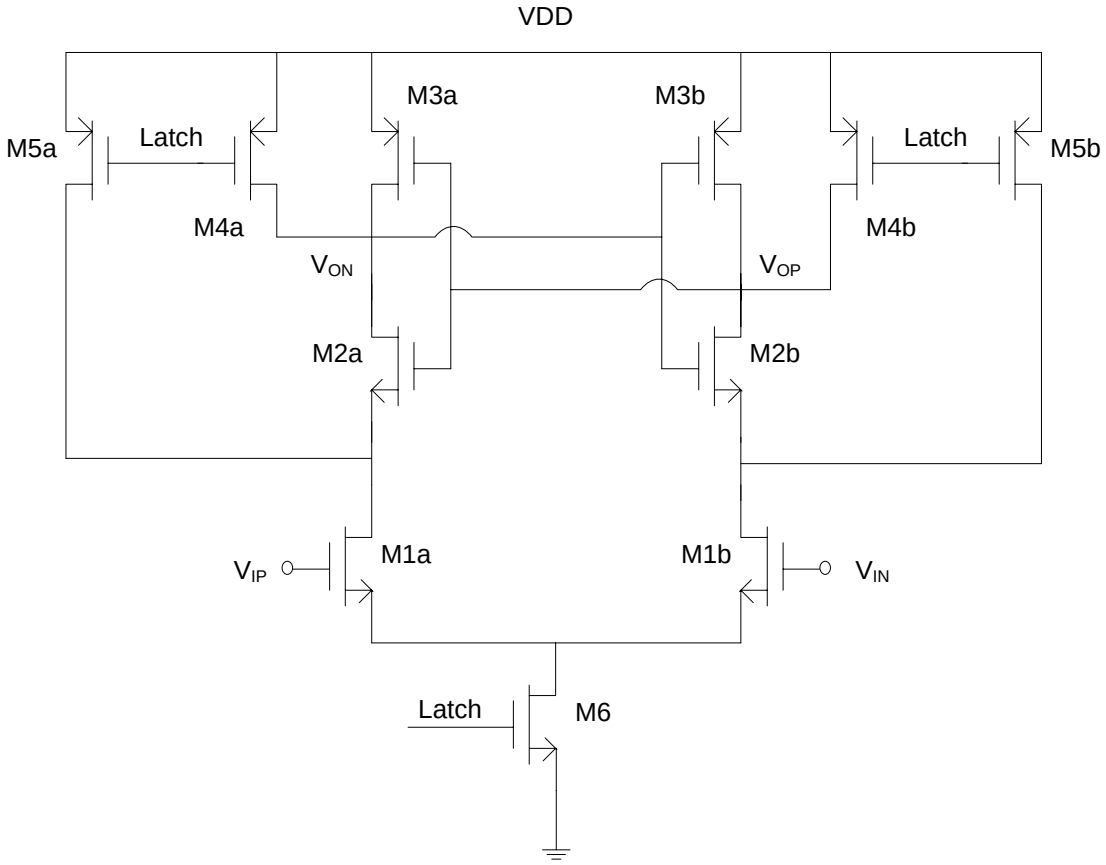


Figure 3.17 A Structure of Dynamic Latched Comparator.

During reset phase, *Latch* = low, transistors M_{4a}/M_{5a} and M_{4b}/M_{5b} are on, and reset the output nodes and the drains of differential pair (M_{1a}/M_{1b}) to VDD [10]. Besides, M_6 is off and no current flow in the circuit.

During regeneration phase, *Latch* = high, transistors M_{4a}/M_{5a} and M_{4b}/M_{5b} are cut off while M_6 turns on and current starts flowing into the differential pair. And the source voltages of M_{2a} and M_{2b} begin to decrease, which results in turning on M_{2a} and M_{2b}. But depending on the input voltages, one of the two transistors will turn on first. Thus, the two cross-coupled inverters, M_{2a}/M_{3a} and M_{2b}/M_{3b}, start working and form a positive

feedback loop. Finally, one of the outputs is pulling up to VDD and the other one is pushing down to GND.

For example, during regeneration phase, if $V_{IP} > V_{IN}$, the drain current of M_{1a} is larger than that of M_{1b}, which means the drain voltage of M_{1a} has been pushed down harder than that of M_{1b}, so the drain voltage of M_{1a} is lower. This lower drain voltage results in turning on M_{2a} first and pushing down the drain voltage of M_{2a}, V_{ON}. Very shortly later, M_{1b} will be turned on after M_{1a}, and does the same things, pushing down V_{OP}. However, the faster and first decreasing V_{ON} will turn on M_{3b} and pulling up V_{OP} to VDD whereas increasing V_{OP} further turns on M_{2a} and pushing down V_{ON}. It is obvious that M_{2a}/M_{3a} and M_{2b}/M_{3b} form a positive feedback circuit. Therefore, V_{OP} will finally be pulled up to VDD and V_{ON} will be pushed down to GND.

It has been noted that the dynamic latched comparator hardly consumes static power consumption. The current flow only exists at the beginning of the regeneration phase for a very short period, so it is very power efficient. However, the nodes where the drains of M_{1a}/M_{1b} connect have rail-to-rail excursion, originating a large kickback noise [10]. And there is another type of kickback noise source: the variation of the operating region of the differential pair [10].

Therefore, the dynamic latched comparator is applied because of its fastest speed and most power efficient properties, although it also maintains the largest kickback noise.

3.2.6 Clock

The clock must be applied in a flash ADC system not only to control switched-capacitor circuits but also to coordinate each block functioning correctly. The clock used for switched-capacitor circuits controls the switches on and off in order to sample and hold the signals. And for the whole system, the clock coordinates analog parts and digital parts respectively as well as comprehensively. The principle of how the clock is working in a capacitive interpolated flash ADC will be discussed in the following paragraph. The timing control information of the whole ADC system is illustrated in Figure 3.18. Note that in the thesis, four stages of pre-amplifiers are implemented.

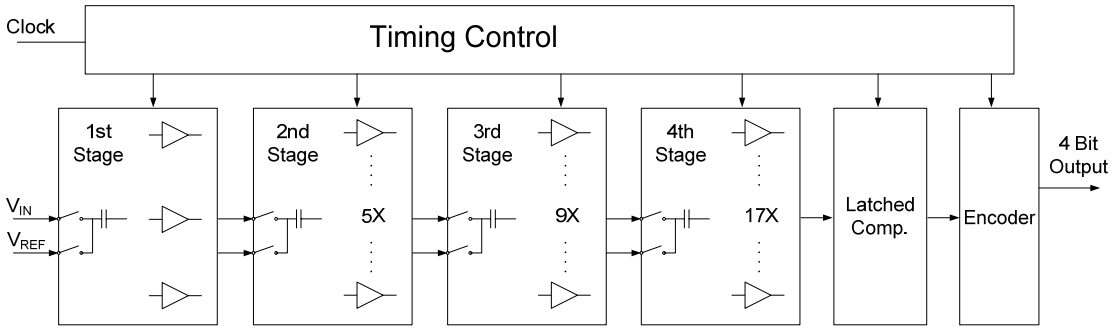


Figure 3.18 Clock Control of the ADC.

The ADC system designed is consists of four pre-amplifier stages, one dynamic latched comparator stage and an encoder stage, all of which must be under control of the right timing system. Each pre-amplifier stage contains differential pre-amplifiers and coupling capacitors that are controlled by clocked switches. The capacitors are used for sample and hold signals as well as acting as interpolation and averaging. The clock

generator will generate two-phase non-overlapping clock signals to control all the switches to make sure they open and close in the correct order.

For analog part (ADC core), during phase1, all the phase1 switches are on and the input signals are tracked through capacitors; pre-amplifiers are working in reset mode. During phase2, phase1 switches off and phase2 switches on, the reference voltages are added to capacitors and pre-amplifiers are working in amplification mode so that the outputs come out. Note that the signals will pass through the 4 pre-amplifier stages and one dynamic latched comparator stage in sequential by only one clock cycle.

The encoder is also synchronized by the clock signals. A bunch of D flip-flops are used in front-end of the encoder to capture the outputs of analog part (thermometer codes), so the timing must be controlled very carefully to make sure only the wanted signals are captured. More detailed information will be discussed in the following part.

3.2.7 Encoder

Encoder, also called digital backend, is used to convert the thermometer codes into binary digits for DSP processing. A typical structure of an encoder is illustrated below.

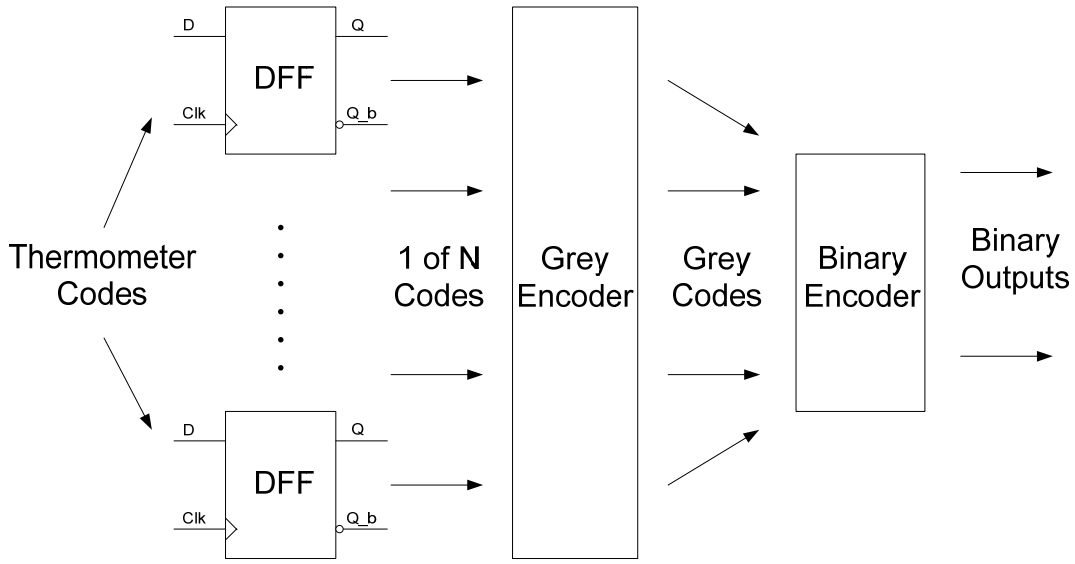


Figure 3.19 A Structure of an Encoder.

The thermometer codes (15 thermometer codes in this thesis) are first captured by D flip-flops resulting in a series of 1-of-n codes, which contains only one “1” and others are “0”. Then these 1-of-n codes are encoded to 4-bit Grey codes. Finally, Grey codes are converted into 4bit binary digits.

Because of metastability and sparkle code issues shown in the previous chapter, Gray code is applied. The good property of Grey code, changing only one digit at each time, endows itself with correcting one error of the thermometer codes (either a metastability or a sparkle code).

The most critical part is the DFF. When thermometer codes come, the clock signal should guarantee that its rising or falling edge must sit in the right location and have enough setup and hold time for triggering. The timing is depicted in Figure 3.20.

Note that DFF is implemented by CMOS construction using only inverters and transmission gates.

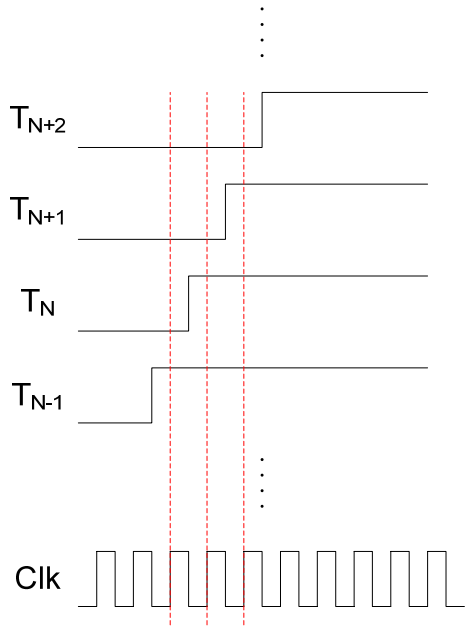


Figure 3.20 The Timing Requirement for DFF to Capture Thermometer Codes.

3.2.8 Low Voltage Differential Signal

For high speed digital systems the traditional CMOS output buffer can not work well to drive the off chip capacitors because of the high voltage swing. In this situation low voltage differential signal (LVDS) is good choice because of the high power efficiency and low power supply voltage LVDS. It can speed up the data transfer frequency up to giga Hz with much lower power dissipation then the traditional emitter coupled logic (ECL). LVDS uses differential data transformation mode with low-voltage

swing (250-400mV). Typically the LVDS signal varies in magnitude from 1.05V to 1.45V at the common mode voltage of 1.25V. The resistance load for LVDS driver is 100 Ω (50Ω for each terminal). So the LVDS driver can be seen as a current source with switched polarity as shown in Figure 3.21.

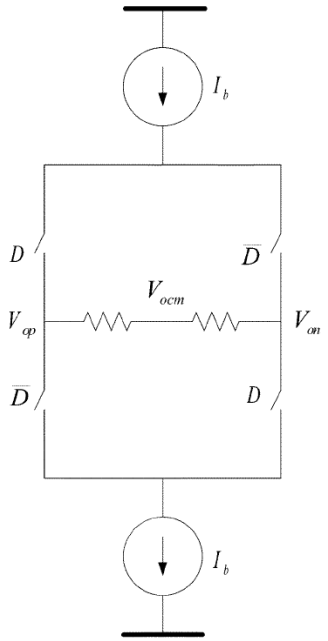


Figure 3.21 Principle of a LVDS.

The four switches are controlled by the digital signal D and its inverse signal $\overline{D}$. When digital signal D is high, the current I_b go through the resistors from node V_{op} to node V_{on}. The voltage from node V_{op} to node V_{on} is determined by the product of the resistors value (100) and the current value. As the resistance load is 100Ω, the bias current determines the output swing. When the digital signal is low, the current go from node V_{on} to node V_{op}. The voltage swing is the same with the previous case but the

polarity is opposite. LVDS receiver can capture the polarity of the differential signal and amplify the signal to rail-to-rail. The recovered signal can be used by another chip or equipment. This is the theory of LVDS driver circuits. In some case when the data rate is very high, we much increase the bias current to drive the off chip capacitance to satisfy the speed requirement.

3.3 Prototype Design

3.3.1 Schematic Design

This 4-bit 2.3GSps capacitive interpolated flash ADC is designed under *SMIC 90nm* technology. It consists of several blocks, including a clock generator, an ADC core (composed of pre-amplifiers and dynamic latched comparators) an encoder and of course a bias circuit. The top schematic of the ADC is illustrated in Figure 3.22.

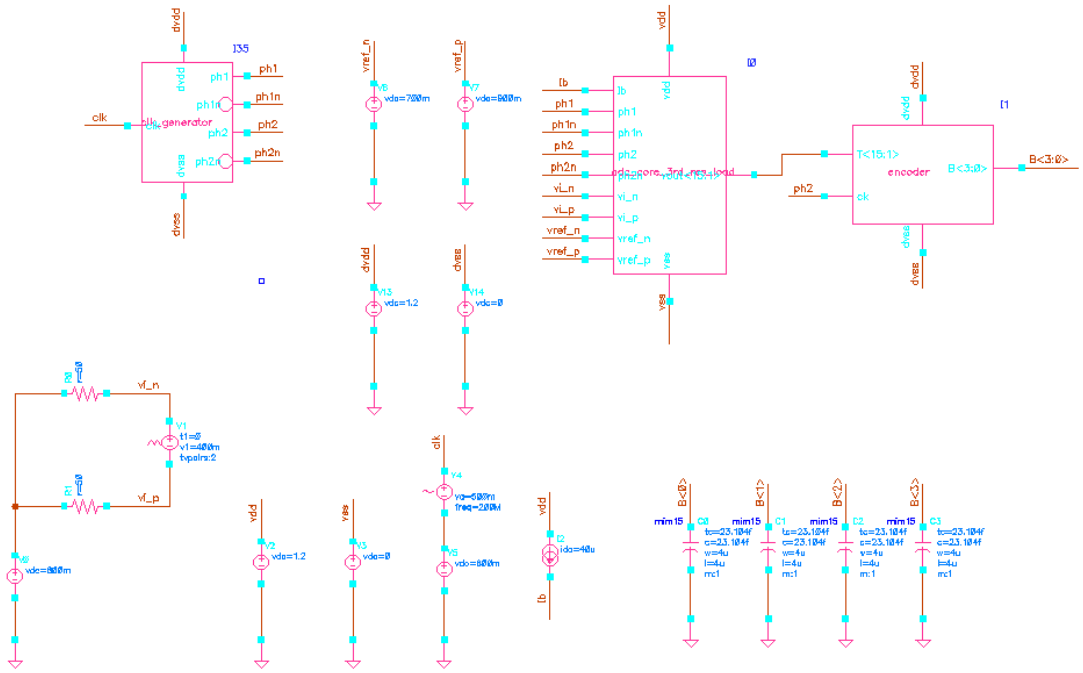


Figure 3.22 Top Schematic of the ADC Block.

The clock generator is used to handle switches in the ADC core system (including pre-amplifiers and dynamic latched comparators) to sample and hold the signal, and as well as the switches in the encoder block.

The clock generator is implemented by two cross-coupled SR latches and a bunch of inverters to increase the driving capability. Two cross-coupled SR latches are used because non-overlap clock signals are required in pre-amplifiers for sampling and holding signals. The schematic is shown in Figure 3.23.

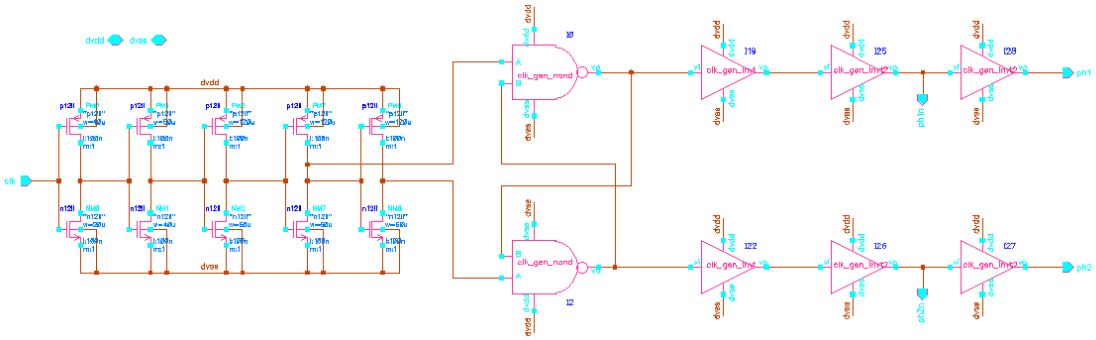


Figure 3.23 A Schematic of the Clock Generator Block.

The ADC comparing block is the core system of an ADC. As discussed before, it consists of pre-amplifier stages, which are used for amplification, and a dynamic latched comparator stage, which is used for determination, output either VDD or GND.

Capacitive interpolation, averaging and offset cancellation techniques are all applied into pre-amplifiers to reduce references, non-linearity and pre-amplifiers' offsets. Two non-overlap clock signals are also used in pre-amplifiers in order to control the switches and let the pre-amplifiers function in proper order. The schematic is shown in Figure 3.24.

Note that the last stage is the dynamic latched comparator stage while the previous four stages are pre-amplifier stages.

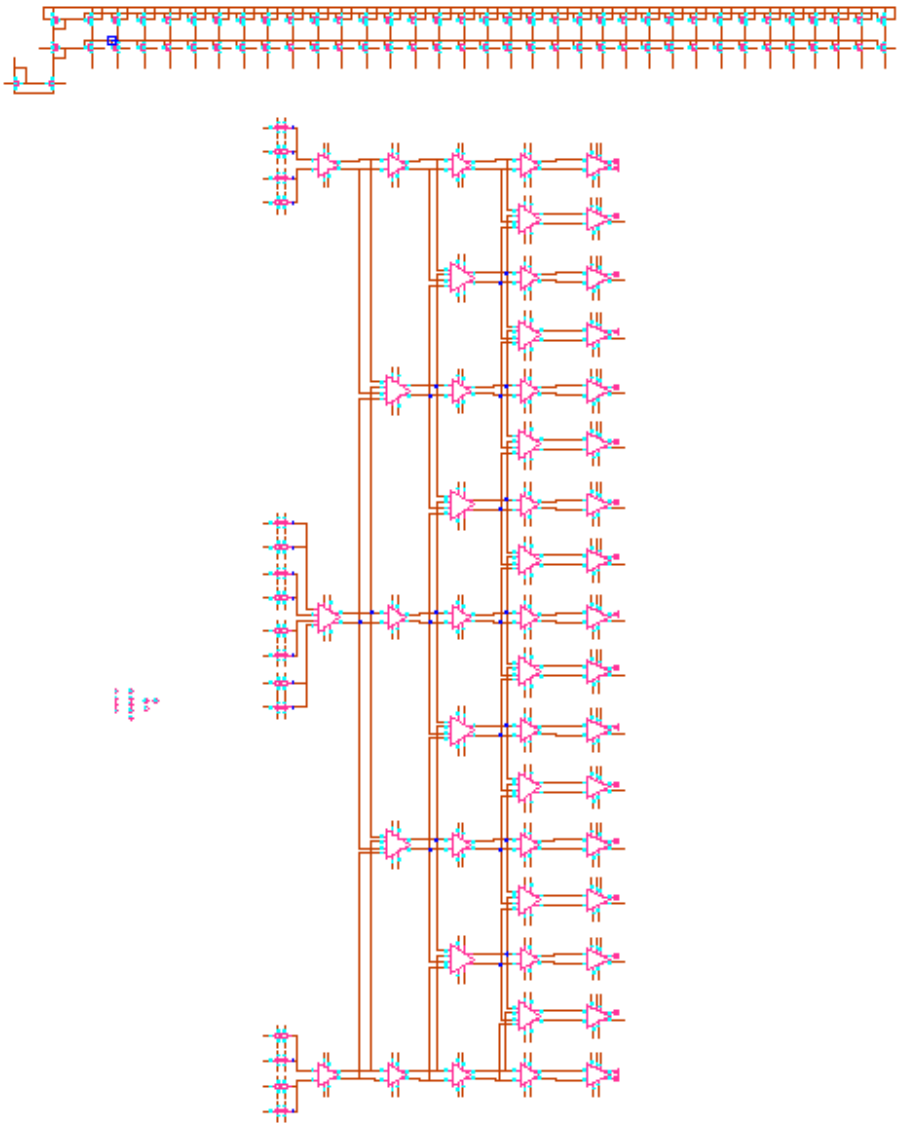


Figure 3.24 A Schematic of the ADC Comparing Block.

From the schematic above, it is noticed that the interpolation factor is 2 for each stage.

In this ADC comparing block, the first four stages are pre-amplifiers, both edge ones and interpolated ones, shown in Figure 3.25 and 3.26 respectively.

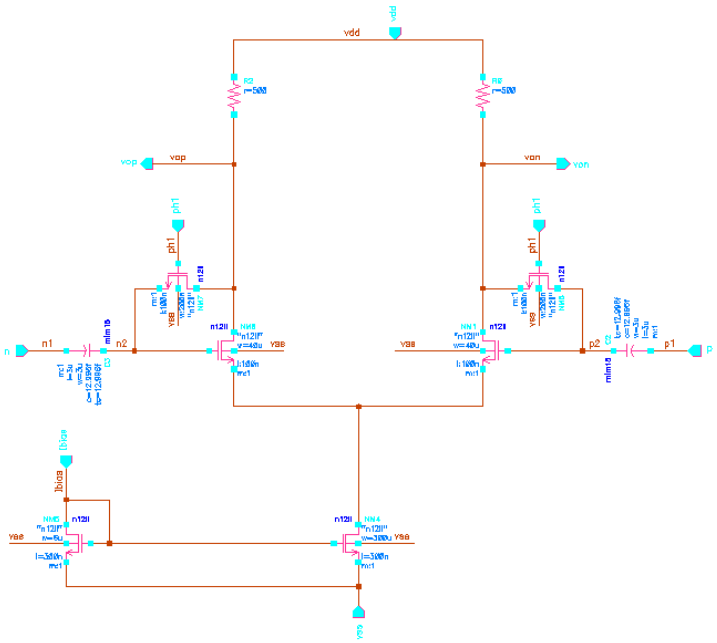


Figure 3.25 A Schematic of the Edge Pre-amplifier.

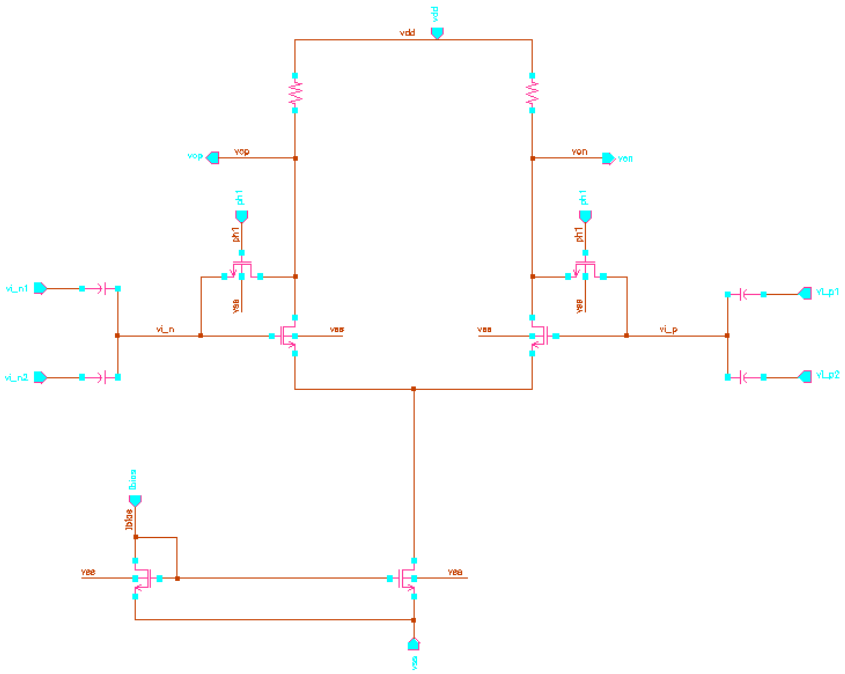


Figure 3.26 A Schematic of the Core Pre-amplifier.

The last stage is called the final comparator, which is composed of a dynamic latched comparator, an SR latch and some inverters. An SR latch is used because after each comparison (output either a VDD or GND) by a dynamic latched comparator the output will be pulled up to VDD (see details in Chapter 3.2.5) again, an SR latch is applied to maintain the comparison result until triggered by next clock (new comparison). The schematic of the final comparator is illustrated in Figure 3.27.

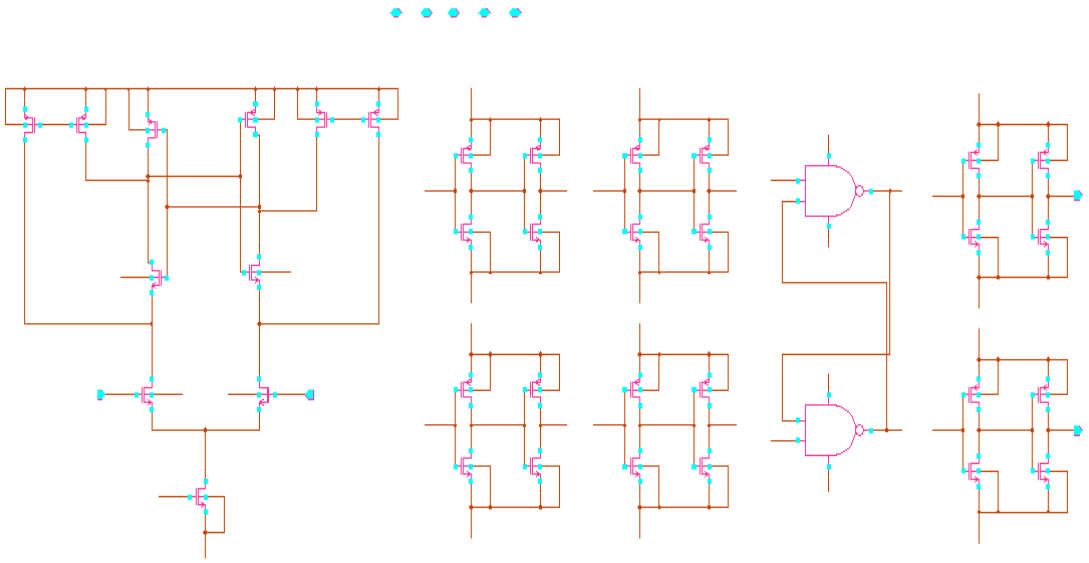


Figure 3.27 A Schematic of the Final Comparator.

The final comparators will output a series of digital thermometer codes. The thermometer codes are supposed to be sent to the encoder block, which consists of some DFFs, NANDs, NORs and inverters, shown in Figure 3.28. The encoder block first converts the thermometer codes into 1-of-N codes, and 1-of-N codes are converted to

Grey codes for preventing the sparkle code and metastability. Finally, the Grey codes will be converted into binary codes as outputs for next DSP processing.

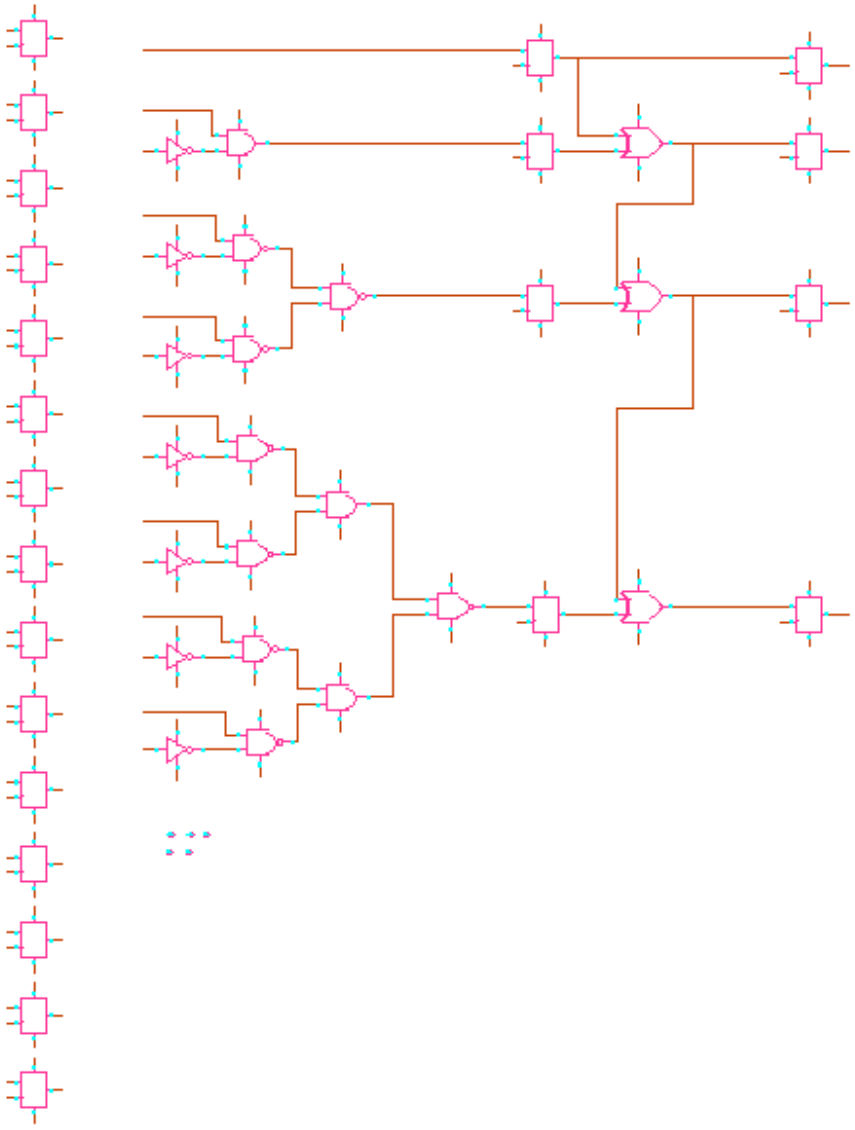


Figure 3.28 A Schematic of the Encoder Block.

The binary codes are sent to the buffer to obtain CMOS outputs or LVDS outputs depending on the practical use. Figure 3.29 shows the schematic of the buffer block.

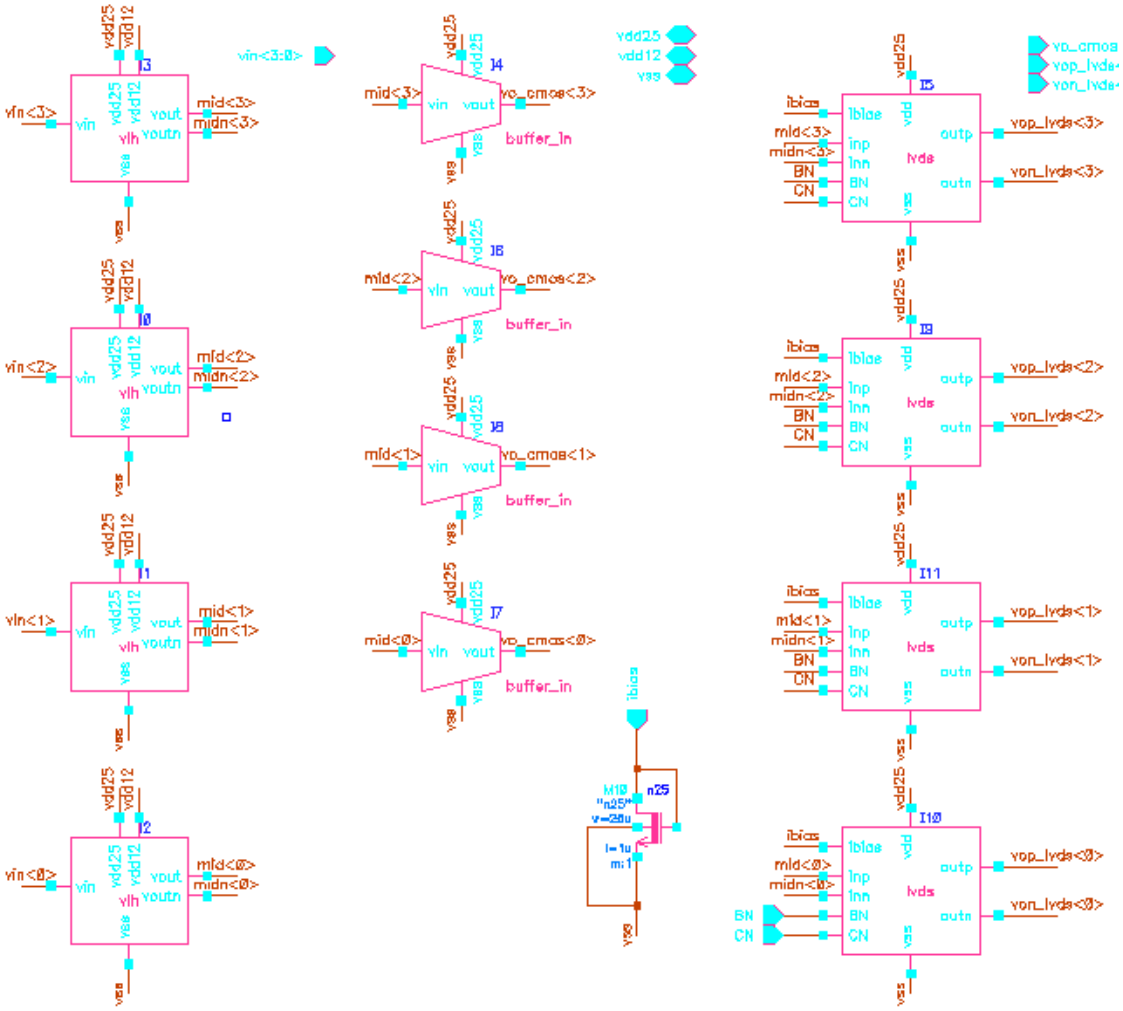


Figure 3.29 A Schematic of the Buffer Block.

Finally, the outputs of the ADC are obtained from the buffer if the ADC is working at high sampling speed or from the encoder directly if it is working at low sampling speed.

3.3.2 Layout Design

Layout demonstrates the physical implementation of an integrated circuit. The process of constructing layouts for mixed-signal integrated circuits has stubbornly defied all attempts at automation. The shape and placement of every polygon and lead require a thorough understanding of the principles of device physics, semiconductor fabrication and circuit theory [10].

Illustrated below is the layout of an improved 4-bit 1GSps capacitive flash ADC, where the original one is from [11].

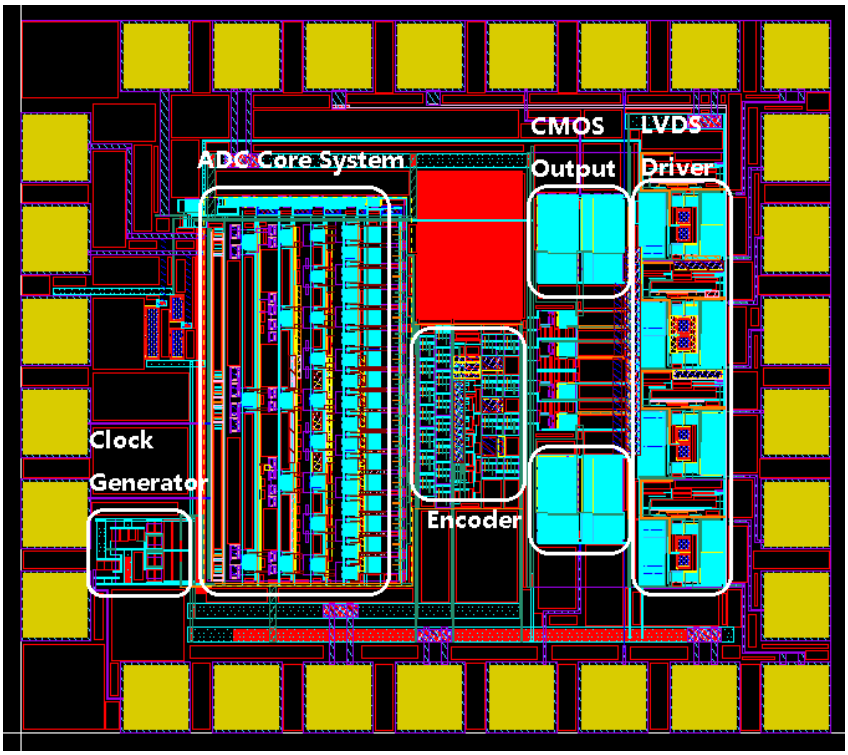


Figure 3.30 Top Layout of the 4-bit 1GSps Flash ADC.

The clock signals used for most other blocks are very critical and should be placed and routed very carefully to avoid any frequency decrease and asynchronization. The encoder is supposed to be separated from the analog ADC core part and then CMOS driver and LVDS driver using I/O devices are placed close to the pads. The LVDS driver is used when the sampling frequency is high. Because of the CMOS characteristic of infinity input impedance the ADC can also work at very low frequency. If this is the case CMOS driver is chosen to save power.

3.4 Simulation Results

Simulation results are shown in this part. Figure 3.31 illustrates the simulation result of the clock generator, where two non-overlap clock signals are generated.

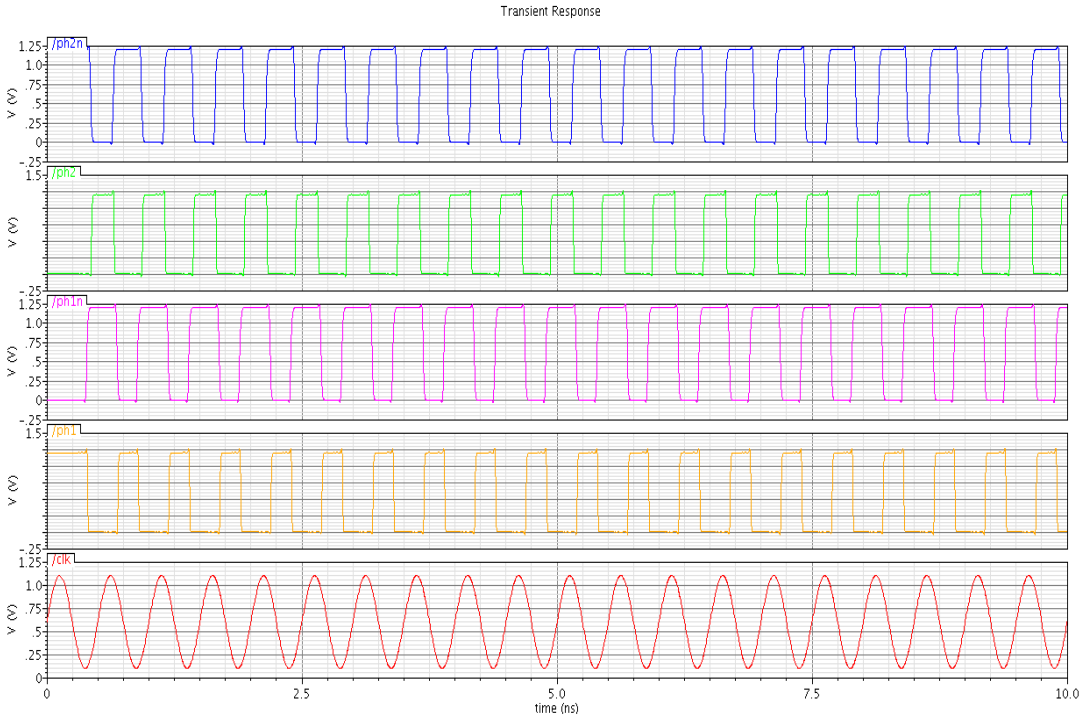


Figure 3.31 Simulation Result of Clock Generator.

From the result, it is noted that a sinusoidal input signal is given to the clock generator, and two phase clock outputs are generated after this block. This is input signal is working at 2.3G Hz.

The final simulation results are shown in Figure 3.32. A ramp signal from the minimum input range to the maximum is used as the input and the outputs are the binary codes rising from 0000 to 1111. Based on the transition point we can estimate the performance of the ADC including INL, DNL, offset, gain error etc.

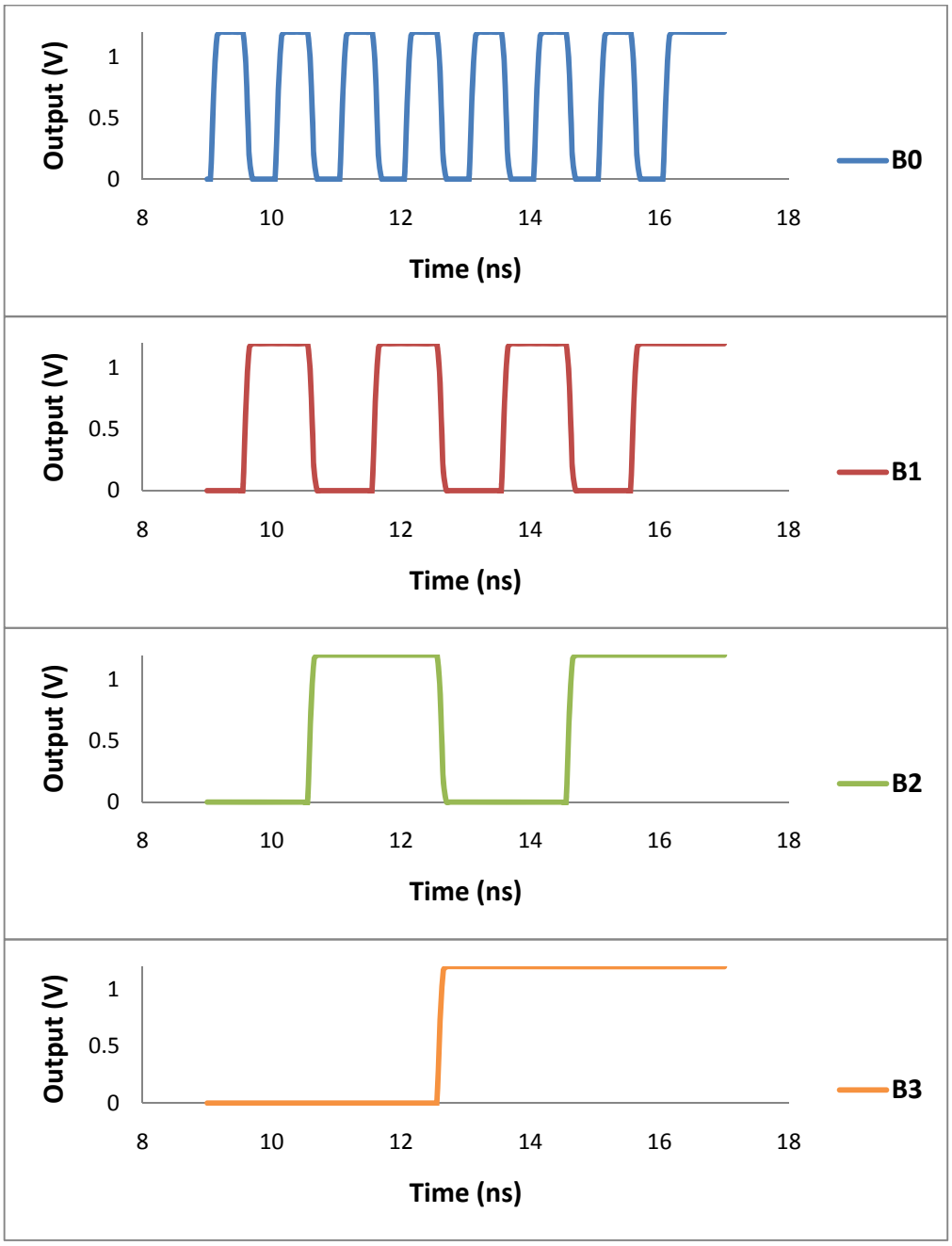


Figure 3.32 Final Simulation Result of the ADC.

3.5 ADC Testing

In this section, typical ADC testing will be presented. As discussed in Chapter 1, a bunch of critical parameters, including static parameters (DNL & INL) and dynamic ones (THD, SFDR & SNDR), are supposed to be tested for ADC circuit. First, a quick review of these parameters is required.

- DNL (differential non-linearity) - the difference between a specified code bin width and the average code bin width, divided by the average code bin width.
- INL (integral non-linearity) - the maximum difference between the ideal and actual code transition levels after correcting for gain and offset.
- THD (total harmonic distortion) - the square root of the sum of squares of a specified set of harmonic distortion components including their alias.
- SFDR (spurious-free dynamic range) - the ratio of the amplitude of the ADC's average output spectral component at the input frequency to the amplitude of the largest harmonic or spurious spectral component observed over the full Nyquist band.
- SNDR (signal-to-noise and distortion ratio) - the ratio of the root-mean-square (rms) amplitude of the ADC output signal to the rms amplitude of the output noise, where noise includes not only random errors but also nonlinear distortion and the effects of sampling time errors.

3.5.1 Sinusoidal Wave Test Set-up

A sine wave test set-up, suggested by IEEE Std. 1241-2000, is supposed to be applied to test the ADC. Sinusoidal waves are chosen for high-speed ADC testing due to the availability of generating high-quality sinusoidal signals, suitability for DFT and dynamic parameter tests, and easy to remove the non-linearity of the signal source [11].

Figure 5.1 illustrates the ADC test set-up environment, which composes of three parts, test signal generator, ADC under test and output data analyzer [11].

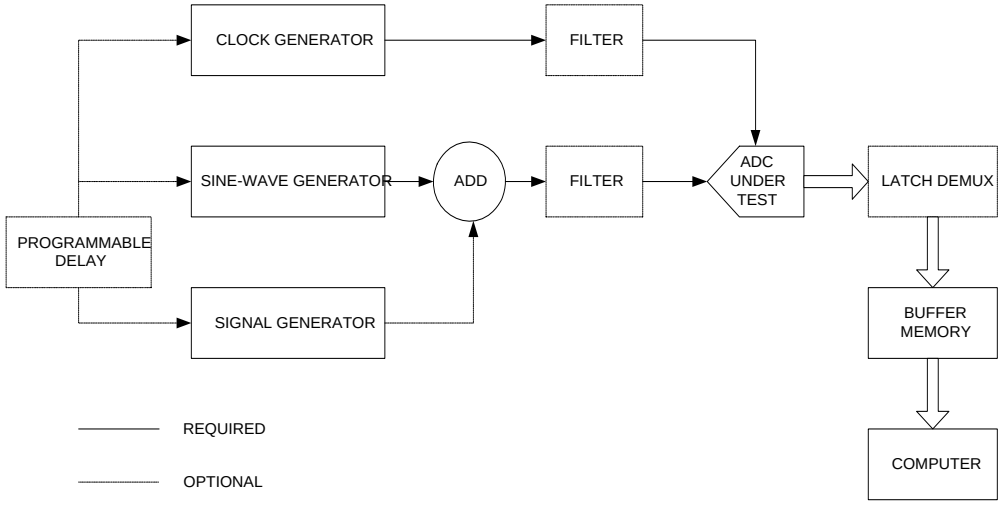


Figure 3.33 An ADC Test Set-up Environment.

The signal generator is producing sinusoidal input signals and clocks. Two/three-tone tests are supposed to be conducted using multiple sinusoidal generators, while a noise generator is applied to provide low-level dither. Noises and harmonics of the input test signal must be reduced by implementing low/band-pass filters. The circuit which is used to capture the digital data samples from ADC is determined largely by the data rate.

Hence, a buffer memory, e.g. a logic analyzer, is recommended to acquire sampled data from a high speed ADC, with the accumulated samples downloaded to a computer at a slower rate [11].

3.5.2 Equivalent-Time Sampling (ETS) Method

Real-time oscilloscopes operating in real-time acquisition mode capture an entire waveform in a single trigger event. This method guarantees a sample rate that is fast enough to get all the samples required to accurately reconstruct the waveform. However, some situations call for higher timing resolution. This can be achieved through an alternative way what is called an equivalent-time sampling and extraction method.

An equivalent time sampling measures only the instantaneous amplitude of the waveform at the sampling instant. In contrast to the real-time sampling, the input signal is only sampled once per trigger. The next time the scope is triggered, a small delay is added and another sample is taken. The intended number of samples determines the resulting number of cycles needed to reproduce the waveform. The measurement bandwidth is determined by the frequency response of the sampler which currently can extend beyond 70 GHz. The equivalent-time sampling is a process where consecutive samples of a repetitive waveform are acquired and assembled from multiple repetitions of the waveform to produce a record of samples representing a single repetition of the waveform [11]. Note that ETS samples the waveform over a number of cycles and can

only be used to measure signals that are repetitive. ETS cannot be used for single-shot or non-repetitive signals. Illustrated in Figure 5.2 is the principle of an ETS.

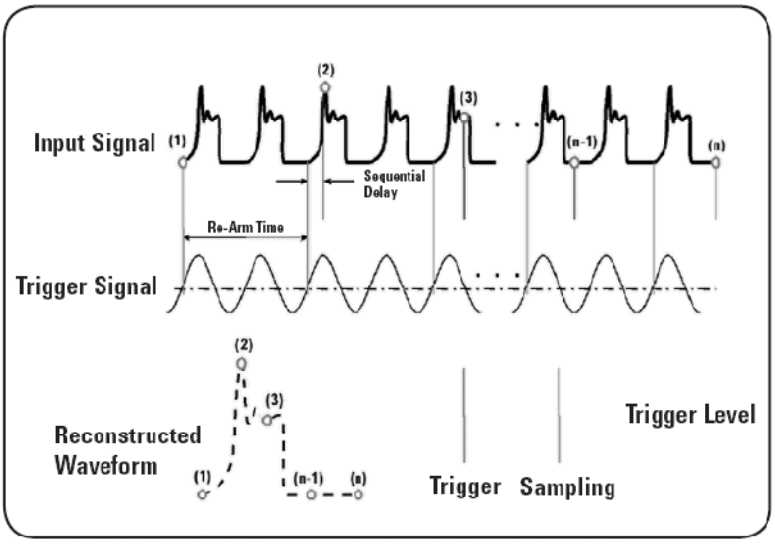


Figure 3.34 Waveforms Acquisition Using an ETS.

The extraction method is carried out by properly choosing repetition rate of an input signal f_t and recording D periods of the input waveform in a single record. After re-arranging the samples with a simple algorithm, a single period of the input signal is obtained as D times the real-time sampling rate. To implement this method, firstly, choose integer D based on the required equivalent sample rate, f_{eq}, such that $f_{eq} = Df_s$, where f_s is the ADC sampling frequency. Secondly, decide the number of real-time samples, L, taken during each repetition of the input waveform by $L = \text{int}(M / D)$, where M is the number of samples in a record. For an ideal ADC transfer characteristic without random noise, the minimum record size M that ensures a sample of every code bin is

$M = \pi 2^N$, where N is resolution of ADC. Finally, input signal frequency f_s, is set to meet $f_i = f_s \frac{D}{LD-1}$, with $LD \le M$. Applying a frequency f_s to guarantee that $LD-1$ distinct sampled phases are uniformly distributed between 0 and 2π radians [11].

3.5.3 Parameter Test Methods

1. Measure DNL & INL: Code transition levels are first determined based on a feedback loop set-up shown in Figure 5.3.

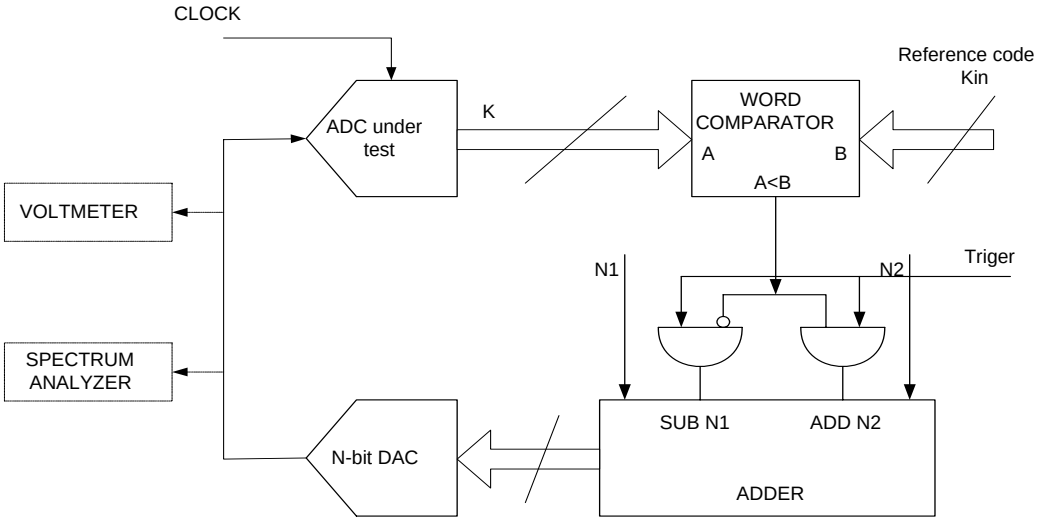


Figure 3.35 A Block Diagram of a Feedback Loop.

Shown above is a so-called digital method distinguishing the previous analog methods. The N-bit DAC generates the feedback signal. N1 and N2 are equal to N0, which is the value that a DAC decreases (A>=B) or increases (A<B) after each conversion cycle when finishing comparing the ADC's output code, A, and a designated reference code, B. This process repeats until the ADC input settles to a stable average

value, which can be either measured by an optional voltmeter or, if the input source is well calibrated, computed from its transfer function. To speed up the iterative process, choose of N0 is critical. During the test, N0 is usually set a little bit larger, and then reduces in binary steps to the point where the appropriate step size is reached for the final settling [11].

With the determined transition level, INL is expressed as a percentage of full scale or in units of LSBs as,

$$INL[k] = 100\% \times \frac{\varepsilon[k]}{2^N \times Q} = 100\% \times \frac{\varepsilon[k]}{V_{FS}} \tag{3.11}$$

where k is the index of the transition level, $INL[k]$ is the INL at output code k, $\varepsilon[k]$ is the difference between $T[k]$ and the ideal value of $T[k]$ computed from G and V_{OS} according to $G \times T[k] + V_{OS} + \varepsilon[k] = Q \times (k-1) + T_1$, G is the gain, nominally equal to unity, V_{OS} is the output offset in units of the input quantity, nominally equal to zero, Q is the ideal code bin width expressed in input units, V_{FS} is the full-scale range of $INL[k]$ for all k and T_1 is the ideal value corresponding to $T[1]$.

DNL is given by $DNL[k] = (W[k] - Q)/Q$, where $W[k]$ is the width of code bin k, $T[k+1] -$ T[k], Q is the ideal code bin width [11].

2. Measure THD, SFDR & SNDR: Total harmonic distortion (THD) are measured by equation below,

$$THD = \frac{1}{M}\sqrt{\sum_{h}\left(X_{avm}\left(f_h\right)\right)^2}$$
$$X_{avm}\left[f_m\right] = \frac{1}{K}\sum_{k=1}^{K}|X_k\left[f_m\right]| \tag{3.12}$$

where m = 0, 1, 2, …, M-1, $X_{avm}\left(f_h\right)$ is the average magnitude of the component at the h^{th} harmonic of ADC output data and M is the number of samples.

Spurious components are the sinusoidal frequency elements other than the fundamental and harmonics, with respect to a full-scale signal. For a pure sine wave input, SFDR is given by,

$$SFDR_{dB} = 20\log_{10}\left(\frac{|X_{avm}\left(f_r\right)|}{\max\left(|X_{avm}\left(f_{sp}\right)| \cup |X_{avm}\left(f_h\right)|\right)}\right) \tag{3.13}$$

The SNDR is the ratio of rms output signal to rms noise, including the distortions from a sine wave input.

$$SNDR = \frac{signal_{rms}}{noise_{rms}} \tag{3.14}$$

3.6 Conclusions

This chapter presents a practical realization of a 4-bit 2.3GSps capacitive interpolated flash ADC.

For this flash ADC, in order to reach high speed, distributed sample-and-hold circuit which is composed of coupling capacitors and pre-amplifiers is applied in the design. The capacitors are also having a function of capacitive interpolation and averaging techniques, where capacitive interpolation network consumes no static power dissipation and capacitance averaging has no edge effect. Offset cancellation is used as well to compress the pre-amplifier's offset.

Final design and simulation results are illustrated in the end.

CHAPTER 4

FLASH ADC DESIGN METHODOLOGY

4.1 Motivation

As mentioned in previous chapters, flash ADCs with interpolation techniques are extraordinary important and have been widely used in high-speed systems. It is generally known that in practical design such high-speed interpolated ADCs design is very changing, very experience-based and trial-and-error work. Unfortunately, there isn't any even qualitative design matrix not mention a quantitative analysis and methodology but only a few [12, 14] discussed some design trade-offs. No wonder flash ADC designers have often been puzzled by complex factors between ADC chip performance and its architecture, circuit, device and technology details.

This kind of experience-based design may work fine before. However, as system performance continues advancing and market demands intensify rapidly, such rough designs are no longer good enough and it is imperative to provide designers a relatively quantitative, accurate, less trial-and-error and more rational design methodology to balance or optimize various design factors. This means a quantitative matrix and clear mapping between the ADC chip specs (e.g., resolution, sampling speed, size and power dissipation) and low level factors, such as, process parameters, device parameters, block circuit parameters, topologies and architectures, is highly required.

Therefore, this thesis was supposed to provide such a quantitative flash ADC design methodology. The thesis researches a comprehensive design matrix analysis and quantitative design approach for capacitive interpolated flash ADCs aiming to address the design challenges. It describes quantitatively the complex relationship among critical factors including ADC speed, interpolation factors, number of stages, pre-amplifier bandwidth, transistor parasitic effects, transistor sizes and technology parameters, etc. The quantitative design technique intends to enable designers to make rapid and predictive decisions in flash ADC designs to achieve both trade-offs and performance optimization in practice.

4.2 Cascaded Multi-Stage Bandwidth Reduction

To fully understand the interpolation effect on flash ADC, it is necessary to first study a cascaded multi-stage network since the interpolated flash ADC system is quite similar to that type of network.

For a one-stage single-pole amplifier, the frequency response can be easily obtained as,

$$H(\omega) = \frac{A_V}{1 + j\,\omega/\omega_L} \tag{4.1}$$

where A_V is open-loop gain, $\omega_L = \frac{1}{R_{out}C_{out}}$, R_{out} and C_{out} are output resistance and capacitance for the amplifier, respectively.

The -3dB frequency of this amplifier is calculated by setting $\left|H\left(\omega\right)\right| = \frac{1}{\sqrt{2}} A_V$, which results in the bandwidth as,

$$\omega_{-3dB} = \omega_L = \frac{1}{R_{out} C_{out}} \tag{4.2}$$

For a multi-stage (*i.e., n* stages) cascaded amplifier network, shown in Figure 4.1,

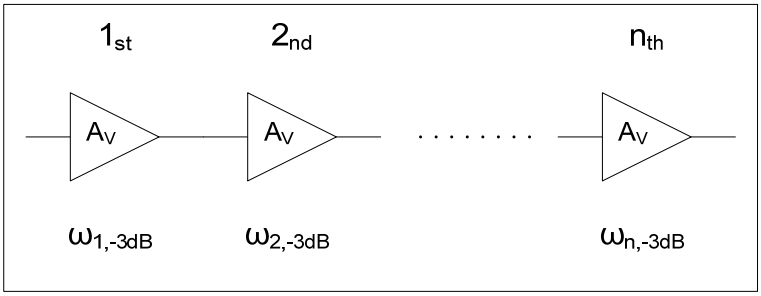


Figure 4.1 A Multi-stage Cascaded Amplifier Network.

assuming all stages are identical to each other, the overall network frequency response is expressed as,

$$H\left(\omega\right) = H_1\left(\omega\right) \cdot H_2\left(\omega\right) \cdots H_n\left(\omega\right) = \left[H_1\left(\omega\right)\right]^n = \left(\frac{A_V}{1 + j\,\omega/\omega_{1,L}}\right)^n \tag{4.3}$$

Let $\left|H(\omega)\right| = \left|\left(\frac{A_V}{1 + j\,\omega/\omega_{1,L}}\right)^n\right| = \frac{1}{\sqrt{2}}A_V{}^n$, therefore the system overall bandwidth is obtained,

$$\omega_{-3dB,total} = \omega_{1,L} \cdot \sqrt{2^{\frac{1}{n}} - 1} = \omega_{1,-3dB} \cdot \sqrt{2^{\frac{1}{n}} - 1} \tag{4.4}$$

From the equation above, it is readily observed that the total bandwidth for a multi-stage cascaded amplifier network is significantly reduced and narrower than that for a single-stage amplifier circuit. This overall frequency reduction effect must be taken into consideration in the interpolated flash ADC designs to optimize the ADC architectures and hence the overall chip performance according to specs.

4.3 Bandwidth Analysis for Capacitive Interpolated Flash ADCs

Generally, an interpolated flash ADC core consists of pre-amplifier stages and the comparator stage. Figure 4.2 illustrates a typical structure with an interpolation factor of 2.

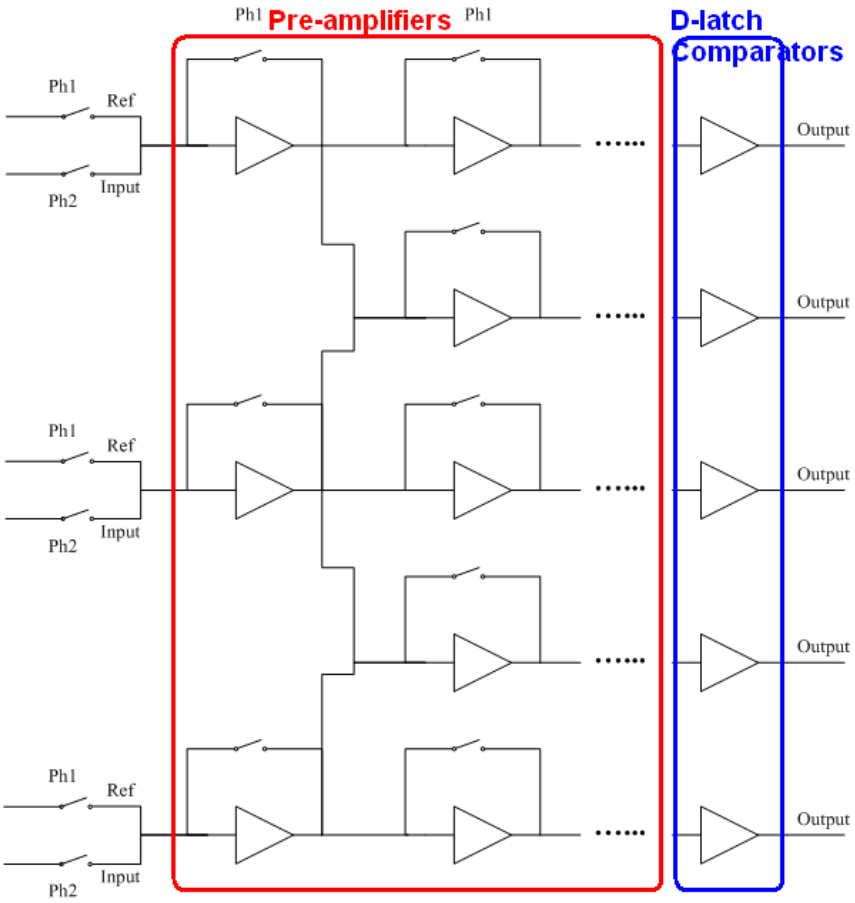


Figure 4.2 A Typical Interpolation Structure.

The number of the pre-amplifier stages varies depending on the ADC bit resolutions and the interpolation factors required. In general, a higher ADC resolution requires more stages. On the other hand, use of a larger interpolation factor helps to reduce the number of stages. Detailed impacts of the interpolation flash ADC structure on the ADC bandwidth will be discussed in this section.

Basically speaking, different signal propagation paths and loading effects have different contributions to the overall system's bandwidth. Figure 4.3 depicts the effects and locations of pre-amplifiers in a practical 4-bit capacitive interpolated flash ADC.

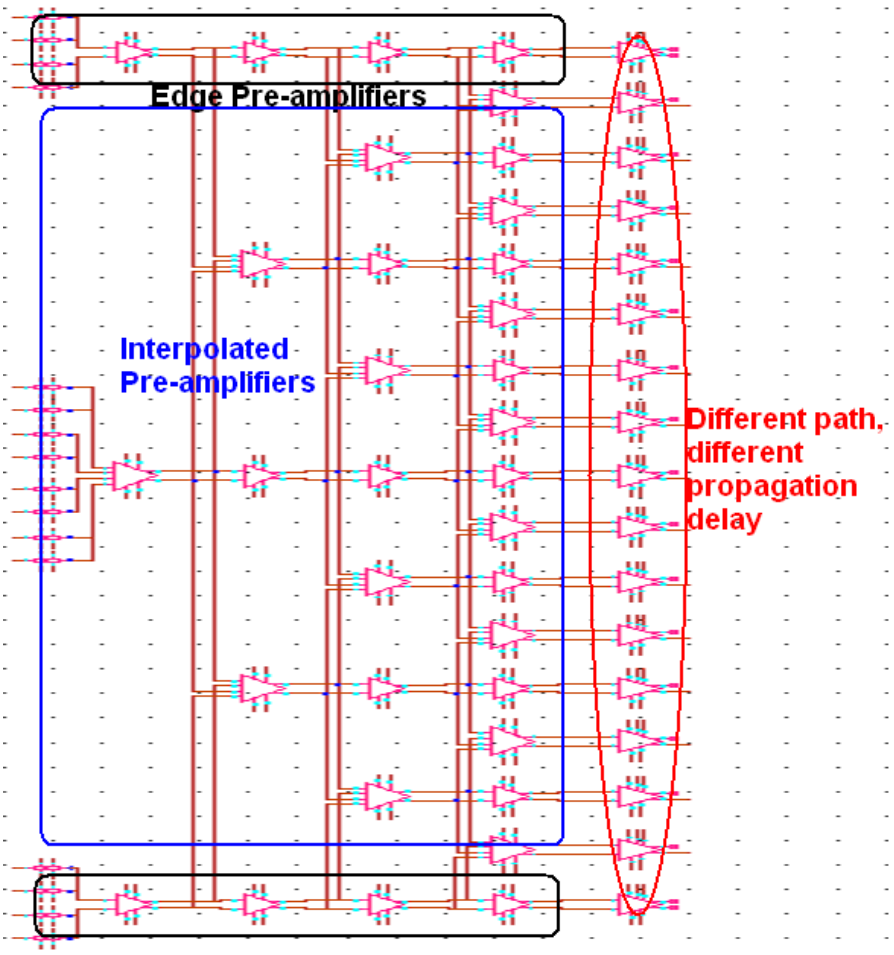


Figure 4.3 Different Located Pre-amps Having Different Effect on Overall Bandwidth.

This interpolated ADC core consists of edge pre-amplifiers and interpolated pre-amplifiers. As discussed before, the cascaded pre-amplifiers, used for interpolation purpose, will reduce the overall bandwidth because of the increase of the number of

stages. And this reduction results in direct decrease of the overall frequency bandwidth for the flash ADC.

On the other hand, it is realized that the pre-amplifiers at different locations have different loading effects of capacitance and resistance. Obviously, each interpolated pre-amplifier is loaded by three succeeding pre-amplifiers (in practical design sampling capacitors will be applied between previous pre-amplifier and succeeding pre-amplifiers), while each edge amplifier has a lighter loads coming from two succeeding pre-amplifiers. Therefore, it is easy to understand that these interpolated pre-amplifiers will definitely have larger time constants and hence narrower bandwidth, which finally result in reducing the ADC speed. Considering the fact that the comparators in the final stages have little impact on flash ADC speed because they are normally some kinds of dynamic latched comparators, which can compare the signals at a speed much faster than pre-amplifiers, the bottle neck of a flash ADC speed is typically set by the heavily-loaded interpolated pre-amplifiers network.

Therefore, the above analysis makes it clear that two main factors have critical impacts on the flash ADC core bandwidth and hence the sampling speed. The first one is the multi-stage cascaded pre-amplifiers and the second one is the significant loading effects of the interpolated pre-amplifiers. Both structural factors will reduce the flash ADC sampling rate inevitably. Obviously, the multi-stage interpolated pre-amplifier chains with largest loads represent the slowest signal propagation path, which hence determines the final speed of the capacitive flash ADC core, as illustrated in Figure 4.4.

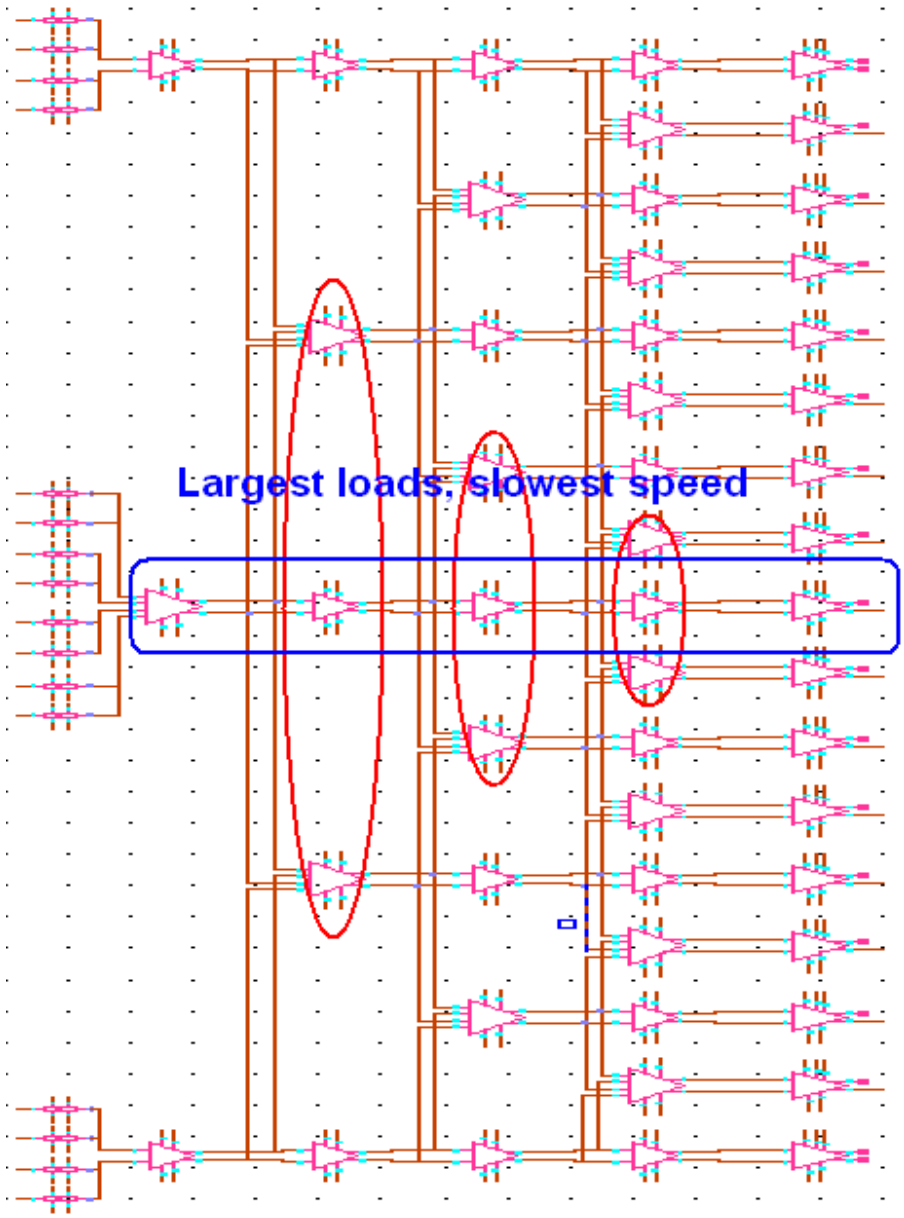


Figure 4.4 One Example of the Slowest Propagation Path.

This can be quantitatively comprehended as following. For an edge amplifier loaded with two succeeding pre-amplifiers, the equivalent load capacitance and resistance are shown in Figure 4.5.

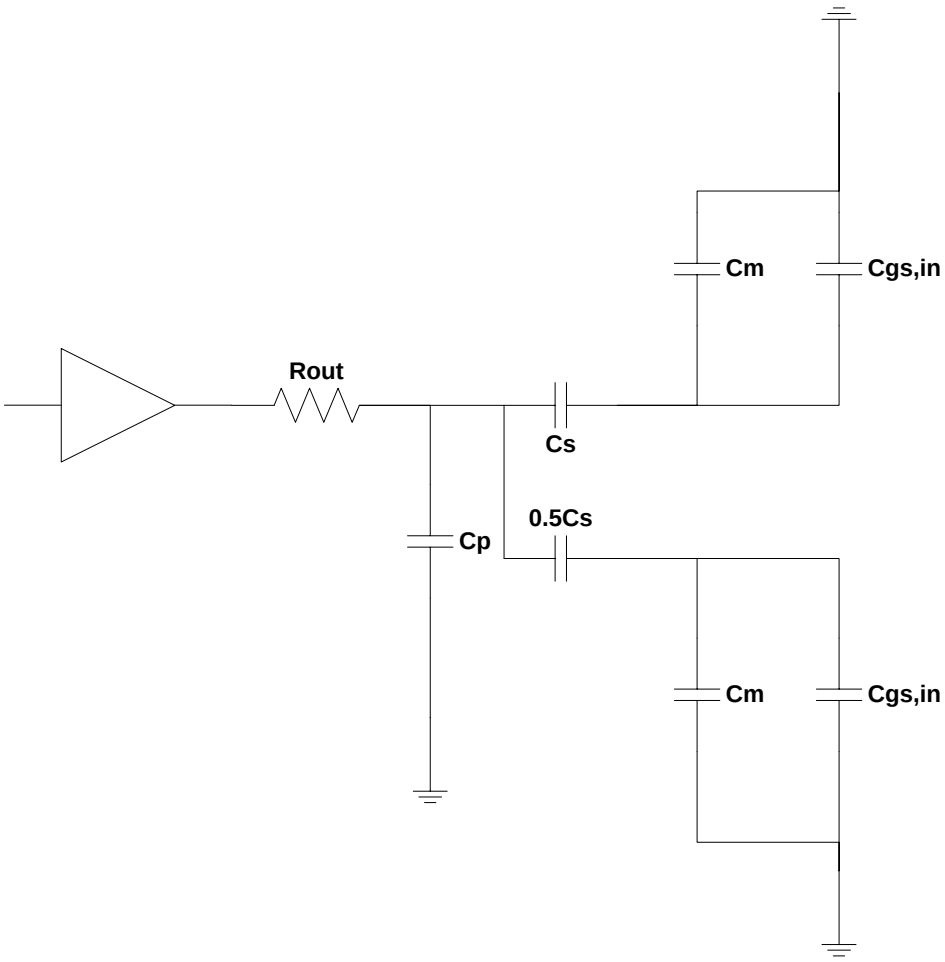


Figure 4.5 An Equivalent Loading Circuit of an Edge Pre-amplifier.

R_{out} and C_p are the output resistance and the parasitic capacitance of the edge pre-amplifier respectively, where C_p includes C_{dg}, C_{ds}, C_{dd}, C_{db}. C_s is the sampling capacitance used in practical design. C_m is the miller capacitance, and $C_{gs,in}$ and $C_{gd,in}$ is the gate-source and gate-drain capacitance of the succeeding pre-amplifier. Therefore the total load capacitance for an edge pre-amplifier is calculated as,

$$\begin{aligned} C_{L,edge} &= C_p + 1.5 \cdot C_s \,//\left(C_m + C_{gs,in}\right) \\ &= \left(C_{dg} + C_{ds} + C_{dd} + C_{db}\right) + 1.5 \cdot C_s \,//\left[\left(1 + A_V\right) \cdot C_{gd,in} + C_{gs,in}\right] \end{aligned} \tag{4.5}$$

Note that one of the sampling capacitance (interpolated capacitance) is only $0.5C_p$, because the capacitor used for interpolation has half the value of normal sampling capacitor. The reason of this is that the voltage after being interpolated reduces to half of the value and in order to maintain the same charging or discharging time the capacitance of the interpolated capacitor is also reduced to half. But it is not a problem when doing analysis if assuming the interpolated capacitance remains C_p for simplicity. From Equation (4.5), the bandwidth of the edge pre-amplifier can be obtained as $\omega_{\text{-3dB, edge}} = 1/\left(R_{out} \cdot C_{L,edge}\right)$. The last stage pre-amplifier may have different loads from the previous ones since it is connected to the comparator. For simplicity, it is reasonable to assume the last stage pre-amplifier still maintain the same load.

However, the interpolated pre-amplifier maintains a larger load capacitance as illustrated in Figure 4.6. The load capacitance can be calculated as below,

$$\begin{aligned} C_{L,central} &= C_p + 2 \cdot C_s \,//\left(C_m + C_{gs,in}\right) \\ &= \left(C_{dg} + C_{ds} + C_{dd} + C_{db}\right) + 2 \cdot C_s \,//\left[\left(1 + A_V\right) \cdot C_{gd,in} + C_{gs,in}\right] \end{aligned} \tag{4.6}$$

It is clear that the load for interpolated pre-amplifier is larger, and hence the bandwidth will be reduced, i.e., $\omega_{\text{-3dB, interpolate}} = 1/\left(R_{out} \cdot C_{L,central}\right) < \omega_{\text{-3dB, edge}}$. Therefore, among all cascaded pre-amplifier networks, the multi-stage interpolated pre-amplifiers

have the lowest bandwidth and the overall bandwidth of the ADC system is mainly determined by the interpolated pre-amplifier network.

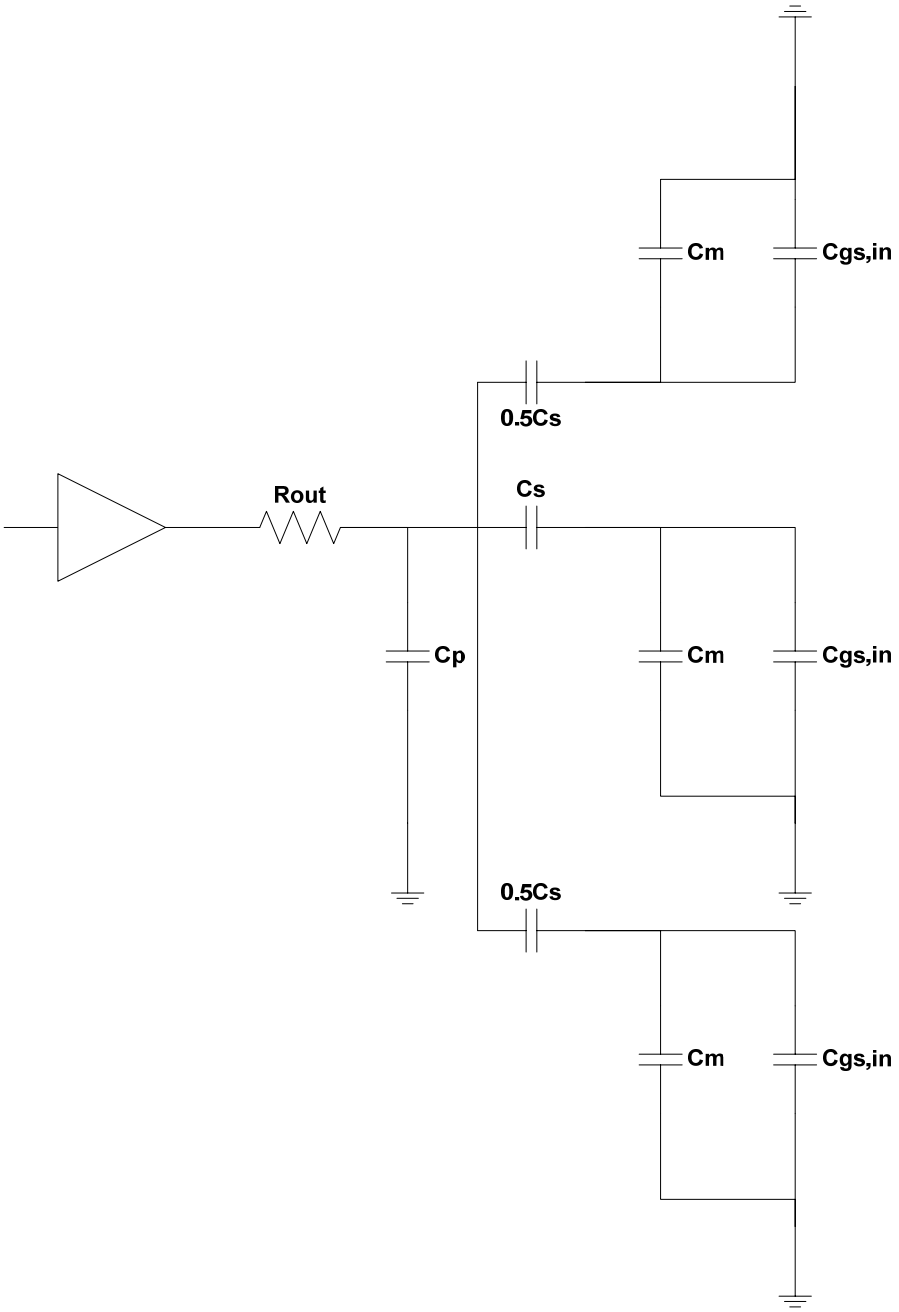


Figure 4.6 An Equivalent Loading Circuit of an Interpolated Pre-amplifier.

Apply Equation (4.6) into Equation (4.4), one can obtain the overall bandwidth of the capacitive interpolated flash ADC as,

$$\begin{aligned}\omega_{-3\text{dB,ADC}} &= \omega_{-3\text{dB,interpolate}} \cdot \sqrt{2^{\frac{1}{n}}-1} \\ &= \frac{1}{R_{out} \cdot \left[C_p + 2 \cdot C_s \,//\, \left(C_m + C_{gs,in} \right) \right]} \cdot \sqrt{2^{\frac{1}{n}}-1}\end{aligned} \tag{4.7}$$

As a result, Equation (4.7) readily illustrates that the overall bandwidth of a capacitive interpolated flash ADC can be substantially reduced by both the multi-stage interpolation technique and the relatively heavier loads of the interpolated pre-amplifiers, and hence the sampling rate is reduced as well.

The previous analysis has clearly illustrated the interpolation's impact on the capacitive flash ADC qualitatively and quantitatively. Furthermore, it also states that whole-chip flash ADC design matrix is desirable and realistic, which maps critical ADC specifications, such as, resolution, sampling rate with key ADC architectural parameters, e.g., interpolation factor and number of stages.

4.4 Sampling Speed Analysis for Interpolated Flash ADC

With the detailed frequency bandwidth analysis for capacitive interpolated flash ADCs discussed previously, this section presents quantitative analysis of sampling rates. Start with a single one-pole amplifier circuit, based upon the previous analysis for the frequency response, its voltage response signal in time domain can be readily obtained as,

$$v_0(t) = A_V \cdot v_{in} \cdot \left(1 - e^{-\frac{t}{\tau}}\right) \tag{4.8}$$

where τ is the time constant given by $\tau = R_L \cdot C_L$, A_v is the low frequency amplifier gain, and R_L and C_L are the load resistance and capacitance of the amplifier, respectively.

It is easy to understand that the time required to amplify the input signals to a certain magnitude value depends strongly on the difference of the input signals. For example, an amplifier is likely to take a longer time to amplify the inputs if the difference of the input signals is smaller; and vice versa.

For an ADC system, the worst case, which means the slowest response, is when the input signal sits in the middle of two adjacent references, which means the difference between the input and either adjacent reference is 1/2 LSB. Figure 4.7 illustrates the worst case and normal case scenarios.

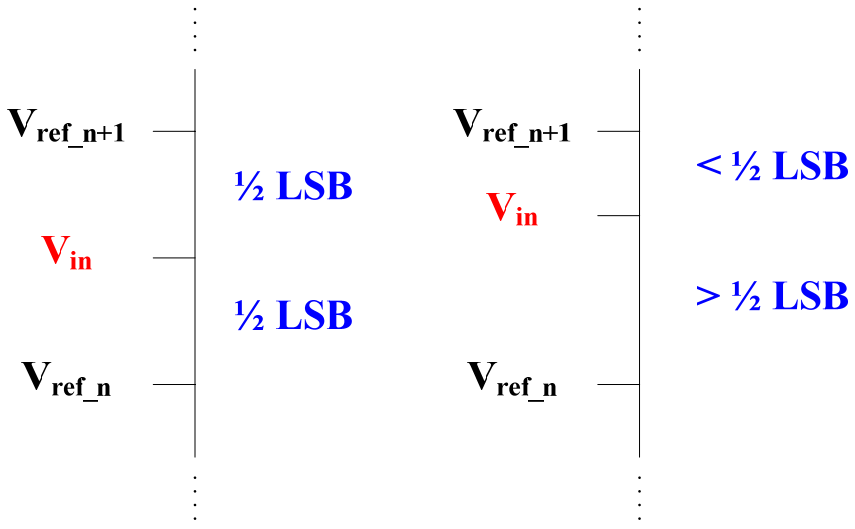


Figure 4.7 Input Sitting Between Any Two Adjacent Reference.

From the figure above, it clearly indicates that if the input is located in the middle of two adjacent references, both $|V_{in} - V_{ref\_n}|$ and $|V_{in} - V_{ref\_n+1}|$ will be equal to $1/2$ LSB; while if the input sits anywhere besides mid-point, one of the differences ($|V_{in} - V_{ref\_n+1}|$) is less than $1/2$ LSB and the other ($|V_{in} - V_{ref\_n}|$) is larger than $1/2$ LSB. Therefore, think about this way, two adjacent multi-stage pre-amplifier networks will both take a time period of t_0 for the amplification if the input sits in mid-point; on the other hand, one multi-stage pre-amplifier network will take a shorter time period t_1 ($t_1 < t_0$) and the other will take a longer time period t_2 ($t_2 > t_0$) if the input sits in a non-mid-point. The latter situation is better for an ADC system if the Grey encoder is applied. Because in the first case both adjacent analog outputs could experience in metastability or even errors but in the latter case only one analog output (smaller difference one) is experiencing in metastability or an error. By implementing Grey encoder, which is immunity to one error, that single error in the latter case could be corrected and all analog outputs would be encoded correctly.

The outputs of pre-amplifier networks will be sent to latched comparators, where output signals should be large enough (larger than a certain minimum value) in order to be sensed by the latched comparators for accurate signal comparisons. If considering the whole multi-stage pre-amplifier networks as one big amplifier system which has only a single pole, one can get a similar equation as deduced in Equation 4.8.

$$v_o(t) = A_{V,overall} \cdot v_{in} \cdot \left(1 - e^{-\frac{t}{\tau_{overall}}}\right) \tag{4.9}$$

where $v_o(t)$ is the output of this one big amplifier, $A_{V,overall}$ is the overall gain of the amplifier and $\tau_{overall}$ is of course the overall time constant.

As it is obvious, $A_{V,overall}$ and $\tau_{overall}$ could be in different values depending on various multi-stage pre-amplifier networks. The largest $\tau_{overall}$ is when the signal propagates through every interpolated pre-amplifier stage as analyzed before. So given a certain time period and an input signal, different multi-stage pre-amplifier networks will result in different values of outputs. In another way saying, if the values of outputs are fixed, smaller gain and input are leading to a slower response time *t*. Thus applying the worst case idea, the following equation can be obtained,

$$\frac{1}{2^{n-1}} \cdot A_V^n \cdot \frac{1}{2} \cdot LSB \cdot \left(1 - e^{-\frac{t}{\tau}}\right) = \Delta v_L \tag{4.10}$$

where Δv_L is the minimum voltage a latched comparator can sense, *n* is the number of pre-amplifier stages, $LSB = \frac{\text{Full Scale}}{2^n}$, τ is the overall pre-amplifier time constant, A_V is the gain of each pre-amplifier, and $\frac{1}{2^{n-1}} \cdot A_V^n$ is the smallest gain that amplifies the input signal, since each interpolated signal from the second stage is only amplified by $\frac{1}{2} \cdot A_V$. Solving Equation (4.10), the slowest response time is easily reached,

$$t_s = \tau \cdot \ln\left(\frac{1}{1-\frac{\Delta v_L \cdot 2^n}{A_V^n \cdot LSB}}\right) = \tau \cdot \ln\left(\frac{1}{1-\frac{\Delta v_L \cdot 2^n \cdot 2^n}{A_V^n \cdot \mathrm{FS}}}\right) \tag{4.11}$$

From Equation (4.11), several things could be concluded:

1.The larger a Δv_L , the longer the response time;

2. The larger an A_V , the shorter the response time;

3. The larger an *LSB*, the shorter the response time;

4. Also it is noticed that larger *FS* or smaller number of bit resolution, shorter response time.

All the observations listed above are making sense. Some of them are really accordance with previous bandwidth conclusions. If Δv_L is larger, which means the latched comparator has a lower sensitivity, the output of the pre-amplifier has to be larger to trigger the latched comparator. Consequently, this larger output will require the pre-amplifiers to take a longer time to amplify the input signal. Similarly, a larger A_V will take a shorter time to amplify the input signal for a latched comparator. However, a larger A_V will certainly lead to a smaller frequency bandwidth. Furthermore, with a larger *LSB*, even if in the worst case, input signal sitting right in the middle of adjacent references,

the response time is faster. Finally, obviously to understand, smaller bit resolution is definitely resulting in faster speed.

Therefore, it is very important to realize that the response time varies according to the design details, e.g., values of pre-amplifier gain, sensitivity of the latched comparator, least significant bit (*LSB*) or full scale range and the bit resolution.

The input signal to an ADC core system is first sampled and then amplified throughout the pre-amplifier chain network. Assume a usual clock duty circle of 50%, amplification will work for only half of the clock period, i.e., $t_s = \frac{T_s}{2} = \frac{1}{2 \cdot f_s}$, where t_s is the response time of the overall pre-amplifier circuit chain, T_s is the clock period and f_s is generally defined as sampling speed for the ADC. Thus, combining previous equations and conclusions, the sampling rate for a flash ADC is,

$$
\begin{aligned}
f_{\text{speed,ADC}} &= \frac{1}{2 \cdot t_s} = \frac{1}{2 \times \tau \cdot \ln\left(\frac{1}{1 - \frac{\Delta v_L \cdot 2^n \cdot 2^n}{A_V^n \cdot \text{FS}}}\right)} \\
&= \frac{\pi}{\ln\left(\frac{1}{1 - \frac{\Delta v_L \cdot 2^n \cdot 2^n}{A_V^n \cdot \text{FS}}}\right)} \cdot f_{-3\text{dB,all}} \\
&= \frac{\pi}{\ln\left(\frac{1}{1 - \frac{\Delta v_L \cdot 2^n \cdot 2^n}{A_V^n \cdot \text{FS}}}\right)} \cdot \sqrt{2^{\frac{1}{n}} - 1} \cdot f_{-3\text{dB,interpolate}}
\end{aligned}
\tag{4.12}
$$

where $f_{\text{speed,ADC}}$ is the overall ADC sampling speed, $f_{\text{-3dB,all}}$ is the overall bandwidth of the cascaded interpolated pre-amplifier chain network, and $f_{\text{-3dB,central}}$ is the frequency bandwidth of each individual interpolated pre-amplifier.

Therefore, the sampling speed of an ADC system is quantitatively established. This is very important, meaningful and exciting conclusion. Equation (4.12) depicts that the sampling speed for a capacitive interpolated flash ADC is quantitatively related with its internal interpolated pre-amplifier bandwidth, interpolation factor, bit resolution and the number of stages used. By developing this quantitative equation, one intends to provide a useful design mapping matrix and methodology for practical IC designers to efficiently design and optimize a flash ADC chip, which are largely experience-based and qualitative oriented so far. Obviously, the sampling speed of a capacitive interpolated

flash ADC is significantly limited by many factors and much slower than a pure theoretical flash ADC. This is mainly because that interpolation function increases the numbers of pre-amplifier stages and their loads, which results in substantial reduction in the total ADC bandwidth and increase of signal propagation.

4.5 Device and Technology Factors in the ADC Design Matrix

Further linkage of ADC specifications with the device and technology factors is discussed in this section. An individual pre-amplifier bandwidth, $f_{-3dB} = 1/(2\pi \cdot R_{out} \cdot C_L)$, can be directly linked to device and technology parameters if expanding the expression, such as, dielectric constant and thickness for gate oxide, transconductance (g_m), drain conductance (g_{ds}), transistor channel width and length (W/L), gate overlap capacitance determined by doping density and junction lateral diffusion, etc. For example, the output resistance, R_{out}, is closely related to $1/g_m$ or $1/g_{ds}$ of a transistor in a given technology, where $g_m \propto W/L$ and $g_{ds} \propto 1/L$. C_L is also related to W and L. Thus the overall ADC sampling speed can be directly linked to the transistor device parameters and process factors.

Consider the cut-off frequency (f_T) of a single transistor, which is generally considered as the maximum working frequency of a transistor, it is found that for an NMOS FET transistor, its cut-off frequency is given as,

$$f_T = \frac{1}{2\pi} \frac{g_m}{C_{gs} + C_{gb} + C_{gd}} \tag{4.13}$$

and g_m is proportional to the overdrive voltage, i.e., (V_{gs}-V_{th}), of an MOSFET transistor if it is not in velocity saturation state. Therefore Figure 4.8 presents the relationship between the cut-off frequency and the overdrive voltage for a single NMOS (minimum size for maximum cut-off frequency) transistor simulated using 90nm foundry CMOS process.

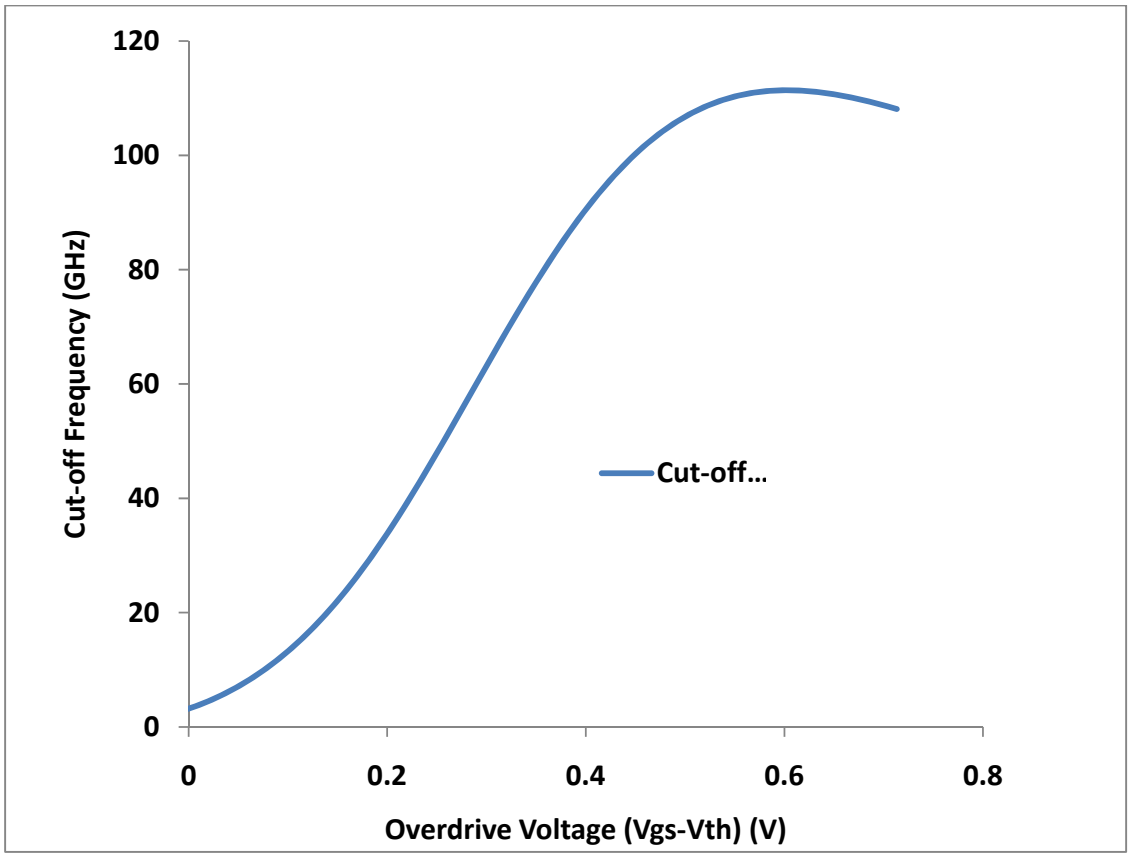


Figure 4.8 The Cut-off Frequency of a minimum-sized NMOS.

Typically, for a 90nm CMOS technology, the overdrive voltage for an MOSFET ranges from 0.1 to 0.3V. Hence, from the figure shown it is reasonable to expect that an MOSFET transistor may easily have a cut-off frequency of approximately 60GHz. If

assuming all the interpolated pre-amplifiers in a capacitive interpolated flash ADC chip are composed of single transistor and neglecting all other possible factors, a theoretically maximum ADC sampling speed can be obtained from Equation (4.14). Taking a flash ADC as an example with interpolation factor of 2, resolution of 4bits, and applying same capacitive interpolation design structure, its overall sampling rate is derived as,

$$f_{speed,ADC} = \frac{\pi}{\ln\left(\frac{1}{1-\frac{\Delta v_L \cdot 2^n \cdot 2^n}{A_V^n \cdot \mathrm{FS}}}\right)} \cdot \frac{2}{3} \cdot \sqrt{2^{\frac{1}{n}} - 1} \cdot f_T \tag{4.14}$$

$$= 2 \cdot \frac{2}{3} \cdot \sqrt{2^{\frac{1}{4}} - 1} \times 70G \approx 0.58 \times 60\mathrm{G} = 34.8GSps$$

Therefore, the maximum sampling speed for this 4-bit capacitive interpolated flash ADC with an interpolation factor of 2 can be theoretically as high as 34.8Gsps. Obviously such highly desired super sampling speed is not practically possible for any ADC in silicon processes at 90nm or even lower technology nodes. Nevertheless, this theoretical value at least manifests how fast an ADC can operate. Furthermore, the quantitative relationship established in this thesis can serve as a valuable design guideline for practical flash ADC designs.

4.6 Design Verification

The quantitative design methodology depicted in previous sections effectively maps the ADC chip specifications with various factors and parameters at architecture levels, block circuit levels, device and technology levels. Experimental validations of this new capacitive interpolated flash ADC design matrix were conducted under two different technologies. The first design example is a 4-bit capacitive interpolated flash ADC in a 90nm foundry CMOS process. The ADC core architecture is the same as Figure 4.3, where the interpolation factor is m=2 and number of stages n=4. Each interpolated pre-amplifier has a frequency response illustrated in Figure 4.9. Note that each interpolated pre-amplifier is loaded by three sampling capacitors and other output parasitic.

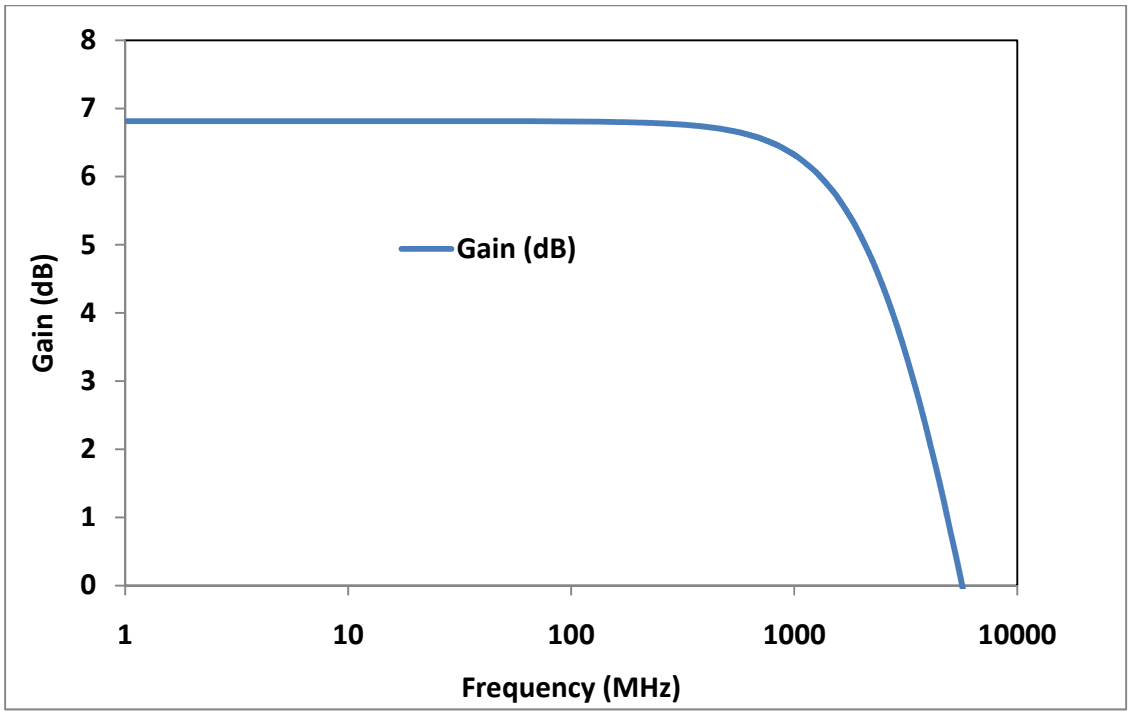


Figure 4.9 Frequency Response of an Interpolated Pre-amplifier in 90nm CMOS.

From the figure, the bandwidth of each interpolated pre-amplifier can be obtained as 3.3GHz. The comparator used in the ADC is a dynamic latched comparator and designed with a minimum sensitivity of 20mV. The gain of each pre-amplifier stage is about two. Apply all above parameters into Equation (4.12), the overall sampling speed for this 4-bit capacitive interpolated flash ADC is obtained as,

$$f_{speed,ADC} = \frac{\pi}{\ln\left(\frac{1}{1-\frac{20mV \cdot 2^4 \cdot 2^4}{2^4 \cdot 400mV}}\right)} \cdot \sqrt{2^{\frac{1}{4}} - 1} \times 3.3 = 2.8GSps \tag{4.15}$$

Equation (4.15) clearly illustrates that the theoretically maximum sampling speed is calculated as 2.8GSps according to the novel mapping methodology. However, the actual simulation result shows that the maximum ADC sampling rate can only reach as high as about 2.3Gsps, which is depicted in Figure 4.10.

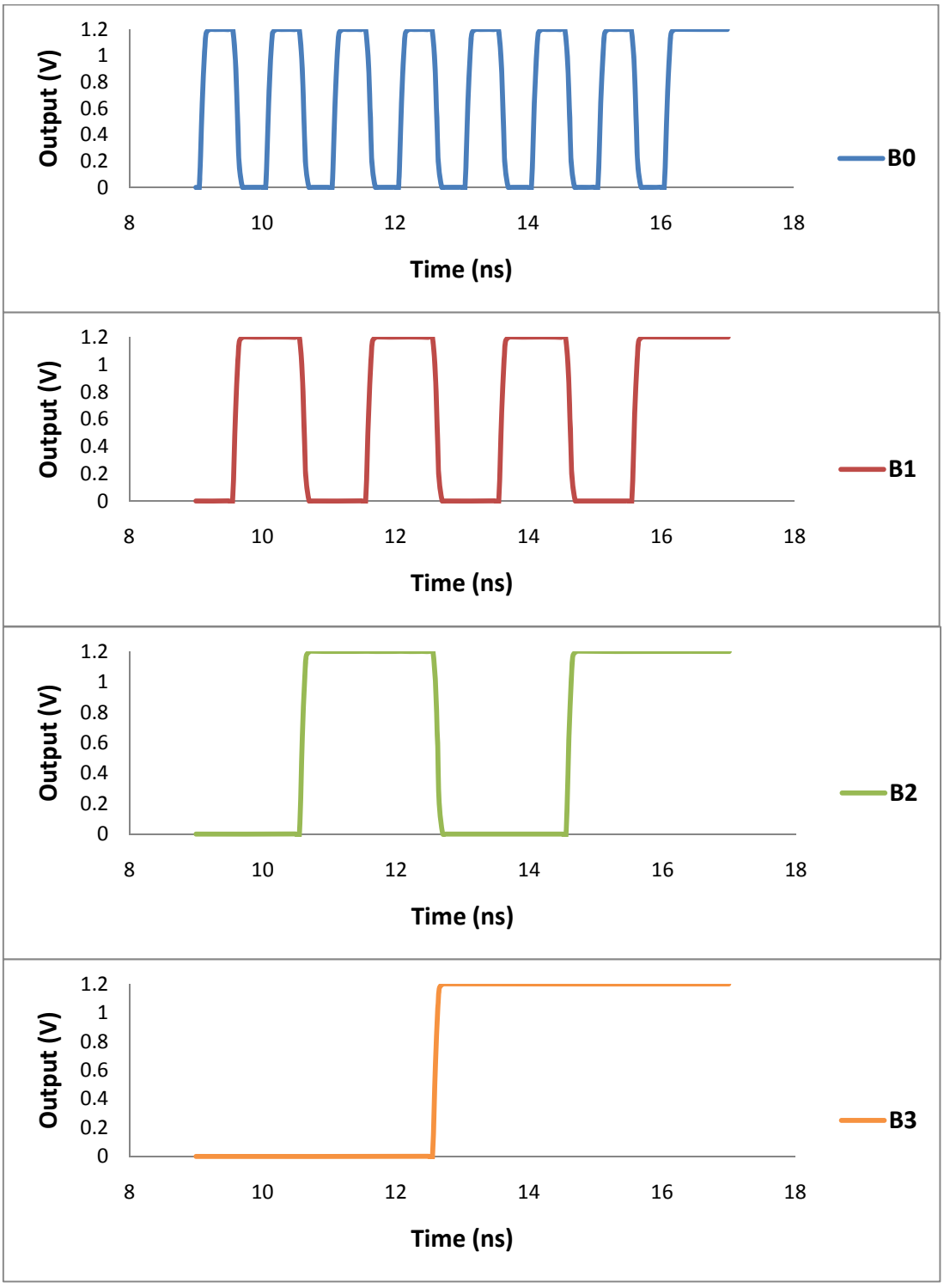


Figure 4.10 4 Bit 2.3GSps ADC Simulation Results in 90nm CMOS.

The figure above shows the final 4-bit 2.3GSps ADC simulation results. Given a ramp input signal, so the outputs are supposed to start at 0000 and increment to the maximum code 1111, which is clearly manifested. This simulation result is reasonable close to theoretical analysis.

To further validate the technology influence on ADC whole chip performance, another example of 4-bit capacitive interpolated flash ADC was designed in a foundry 0.13um CMOS technology. For fair comparison, both designs use the same ADC architecture and topology with an interpolation factor of 2 and other important parameters, such as number of stages, pre-amplifier gain etc. Following the same procedure as discussed before, the individual gain-bandwidth relationship for the interpolated pre-amplifier is obtained as shown in Figure 4.11.

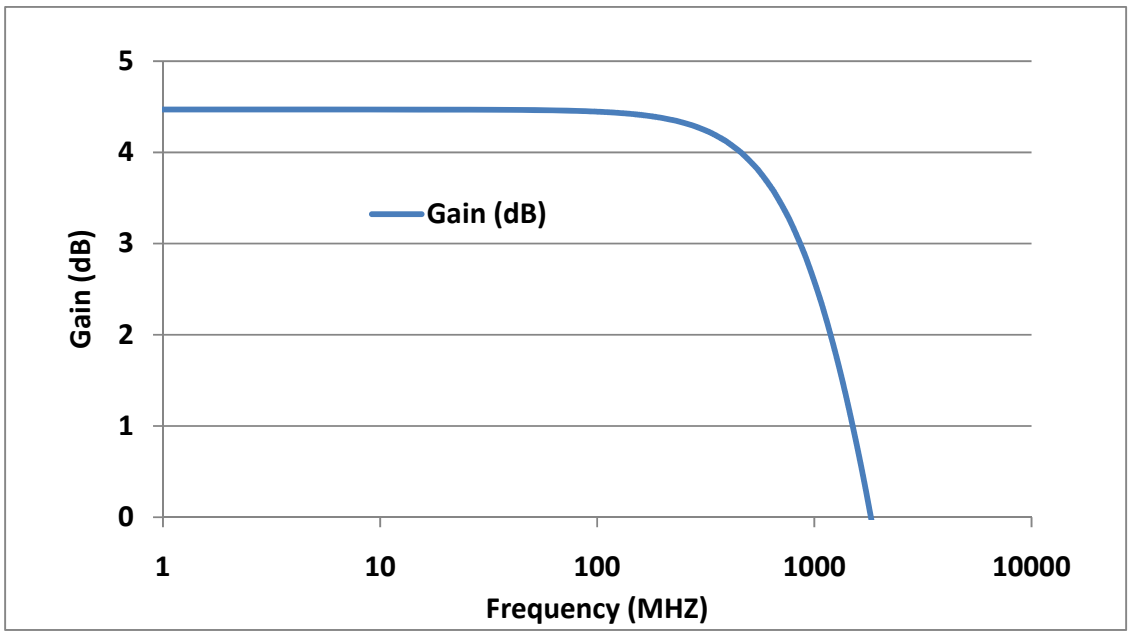
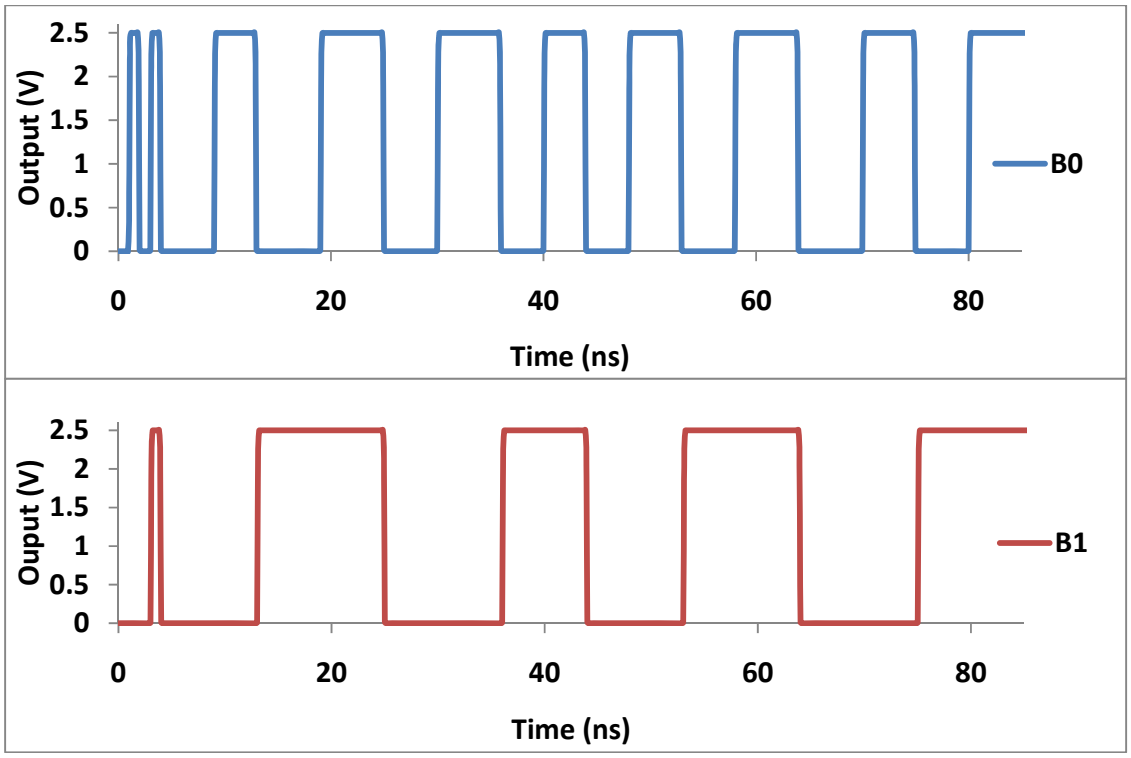


Figure 4.11 Frequency Response of an Interpolated Pre-amplifier in 130nm CMOS.

The bandwidth is obtained as 1.4GHz. Consequently, the final maximum sampling speed for the same flash ADC in the 0.13μm CMOS is found to be,

$$f_{speed,ADC} = \frac{\pi}{\ln\left(\frac{1}{1-\frac{\Delta v_L \cdot 2^n \cdot 2^n}{A_V^n \cdot \mathrm{FS}}}\right)} \cdot \sqrt{2^{\frac{1}{4}} - 1} \times 1.4 = 1.22GSps \tag{4.16}$$

Equation (4.16) clearly illustrates that the theoretically maximum sampling speed is calculated as 1.22GSps according to the novel mapping methodology. Comparing the theoretical value to the practical design, where the simulation results illustrate that the ADC sampling speed reaches to 1GSps, shown in Figure 4.12.



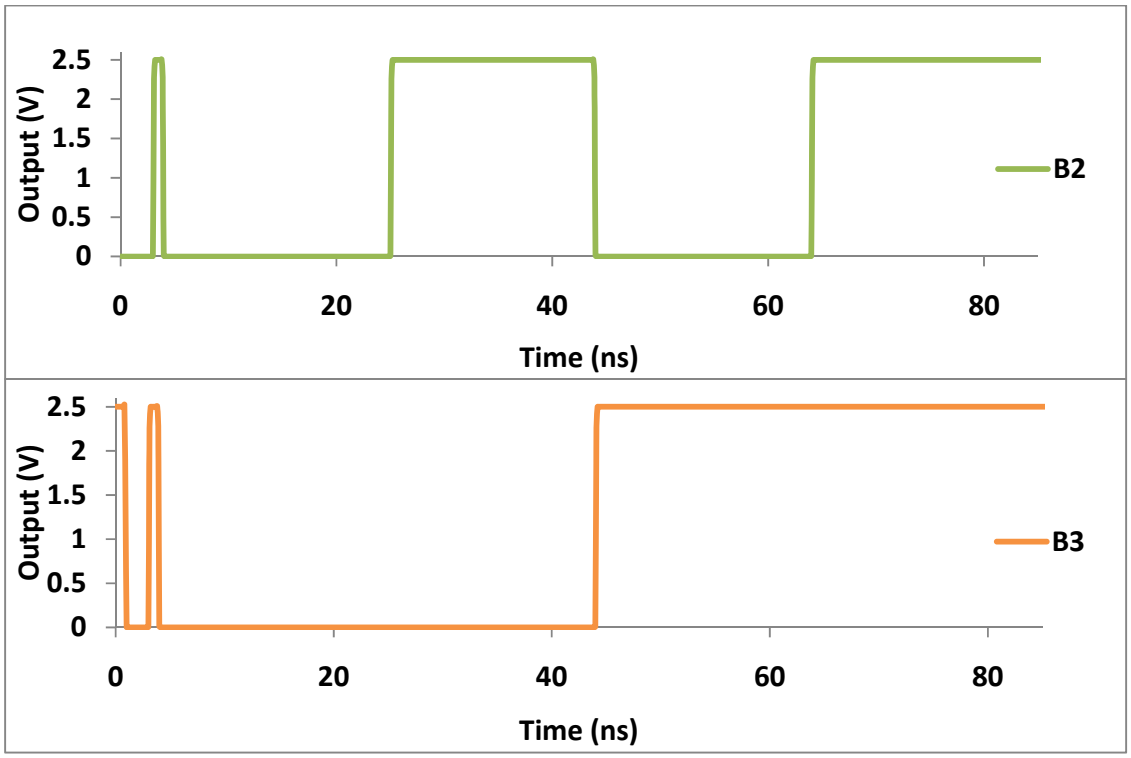


Figure 4.12 4 Bit 1GSps ADC Simulation Results in 130nm CMOS.

The above two capacitive interpolated flash ADC design examples clearly show that the maximum overall sampling speeds from theoretical analysis and simulation results in real-world foundry CMOS PDK files are reasonably close to each other, which strongly supports this new quantitative flash ADC design methodology. On the other hand, if given a set of ADC specifications, the architecture, the critical block circuits, such as, pre-amplifiers, and key ADC structural parameters including number of stages and interpolation factors, can be relatively accurately determined for optimum ADC chip performance using this new ADC design matrix. Although, an offset of about 20% (or even more, which is dependent on technologies) between this quantitative analysis and actual simulation results are observed in ADC designs, which is attributed to the factors

associated with some higher order effects not considered in the simplified equations used, this new quantitative design methodology shall be of significant value in practical capacitive interpolated flash ADC designs. In fact, further improvement of this new quantitative ADC design matrix is on-going currently, taking into consideration of some more factors and doing more accurate circuit analysis.

4.7 Dynamic Power Analysis of Capacitive Interpolated ADC

Power consumption in electronics has become one of the most important concerns in the real world and researchers are trying every method to lower it down. For flash ADCs, power consumption increases dramatically against bit resolutions and sampling speed. As bit resolution goes high in any kind of flash ADCs, the number of pre-amplifiers definitely increases exponentially and hence power dissipation will be extremely large. On the other hand, if the sampling speed continues increasing, design parameters like sizes, currents are supposed to be larger to meet the specs, all of which will cause increase of power dissipation.

Therefore, power estimation of flash ADCs is very much required. For normal flash ADCs, power estimation is easy and straightforward. However, high-speed capacitive interpolated flash ADC power estimation is very complicated due to the capacitors' charging and discharging effects which consume varying but relatively large dynamic power if working at very high speed and with medium-high resolution bits. Current ADC power analysis techniques reported mostly focus on static power only [37,

38]. So, this thesis will provide a rough analysis of power consumption in capacitive interpolated flash ADC.

Consider a switched-capacitor amplifier shown in Figure 4.13. It is obvious that the input capacitors (C_{in} and C_{ip}) are charging and discharging when the switches are turned on and off. Same happens to output capacitors (C_{on} and C_{op}).

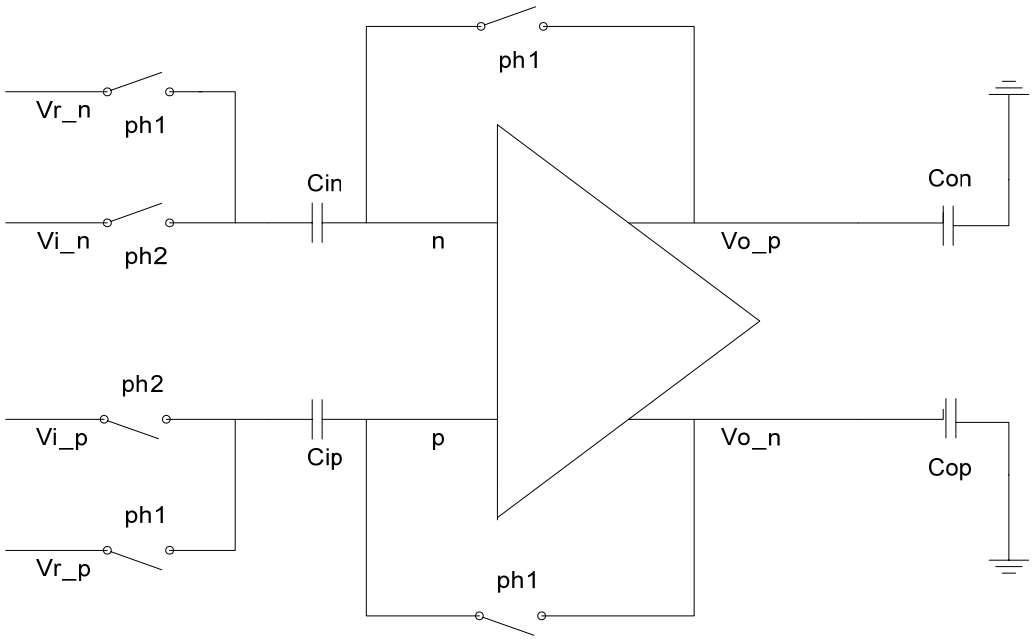


Figure 4.13 A Switched-Capacitor Amplifier.

It is clear that dynamic power will be dissipated during the capacitors' charging and discharging. Calculating the dynamic power is no easy task because of so many uncertainties, such as voltage variations on the capacitors, switching probability, parasitic capacitance and etc. Some simple simulation result shows that the current flowing from the capacitor to ground has periodic spikes when capacitors charge and discharge. This

definitely confirms the consumption of dynamic power. Normally speaking, the dynamic power dissipation can be expressed as,

$$\begin{aligned} P_{D,total} &= P_{\text{in}} + P_{\text{L}} \\ &= \alpha C_{\text{in}} \cdot \Delta V_{\text{in}}^2 f_{\text{in}} + \alpha C_{\text{L}} \cdot \Delta V_{\text{out}}^2 f_{\text{out}} \end{aligned} \tag{4.17}$$

where α is switching probability, C is capacitance including the parasitic one, ΔV is voltage change and f is frequency. Note that the input and output capacitors in capacitive interpolated flash ADC are used for interpolation, hence parasitic capacitances should be considered when calculating dynamic power.

For a practical switched-capacitor pre-amplifier used in an ADC operating at 2GHz, the maximum dynamic power dissipation is estimated as (α=0.5 for simplicity),

$$\begin{aligned} P_{D,total} &= 0.5 \times \left(100\text{f} \times 0.3^2 \times 2\text{G} + 100\text{f} \times 0.6^2 \times 2\text{G}\right) \times 2 \\ &= 90\mu W \end{aligned} \tag{4.18}$$

Here, roughly calculating the input and output load capacitances are about 100fF and the voltage variations of the input and output are 0.3 and 0.6 volts. Comparing the dynamic power with the static power, which is about 2.28mW, the dynamic power is still relatively small for a single pre-amplifier. However, for the whole ADC core, the overall dynamic power could be relatively larger. The capacitive interpolation makes whole chip power estimation complicated, because the edge pre-amplifiers and interpolated pre-amplifiers have different input and output voltages and loading capacitors. For this 4-bit ADC designed in a 90nm technology foundry, the dynamic power is roughly (including 19 edge pre-amplifiers and 15 interpolated pre-amplifiers),

$$
\begin{aligned}
P_{D,overall} &= 19\times 90\mu W + \\
&\quad 0.5\times\left(4\times 100\text{f}\times 0.3^2 + 2\times 150\text{f}\times 0.6^2\right)\times 2\text{G}\times 15 \\
&= 3.87mW
\end{aligned}
\tag{4.19}
$$

Meanwhile, the overall static power consumption is calculated as 77.52mW. Hence, dynamic power is about 5% of the static power. Although still relatively small, the dynamic power does increase faster than static power because of interpolation, which makes some pre-amplifiers have more input capacitors and heavily loaded, resulting in much more dynamic power dissipation.

As sampling speed increases and bit resolution gets higher, the dynamic power can no longer be ignored. For example, at a speed of 10GSps, rough calculation shows that the dynamic power dissipation could easily reach to at least 50mW, which becomes non-negligible. The static power dissipation increases too, reaching to several hundreds of mW. On the other hand, if the resolution bit increases, more stages and more interpolation are needed, which result in more pre-amplifiers and hence certainly larger dynamic power dissipation too.

4.8 Conclusions

This section presents a practical design methodology technique for designing capacitive interpolated flash ADCs, which allows quantitative analysis of influences of technology, device, circuit and structural parameters on whole-chip ADC performance. The quantitative mapping between ADC performance specs and design parameters, such

as, interpolation factor, number of stages, pre-amplifier bandwidth, loading effects, transistor size, technology parameters, etc., will enable IC designers to conduct quick and rational flash ADC design with optimum design balance and whole chip performance. This section also develops a roughly but newly accurate dynamic power analysis technique, which provides dynamic power estimation for capacitive interpolated ADC. This dynamic power dissipation cannot be ignored at high sampling speed or medium bit resolution. Finally, the design methodology was validated by practical design of 4bits flash ADCs in commercial 90nm and 0.13μm CMOS technology.

CHAPTER 5

CONCLUSIONS

5.1 Conclusions

The thesis presents a new comprehensive design methodology for capacitive interpolated flash ADCs, which was verified by design of a 4-bit 2.3GSps flash ADC.

The new ADC design methodology provides a comprehensive and quantitative mapping between chip level ADC performance specs (e.g., sampling speed, bit resolution, etc.) and the critical design parameters at different levels, such as, interpolation factor, number of stages, pre-amplifier bandwidth, loading effects, transistor sizes, technology parameters, etc. This new ADC design methodology aims to provide IC designers with comprehensive, quantitative, yet handy design guidelines to conduct practical flash ADC designs to achieve well-balance and overall optimized ADC chip performance. A new accurate dynamic power analysis technique for capacitive interpolated flash ADC is described, which allows more accurate power estimation including the non-negligible dynamic power dissipation at high sampling speeds for ADC.

The design technique was validated by actual designs of 4-bit flash ADCs in commercial 90nm and 130nm CMOS technologies. Capacitive interpolation, averaging and offset cancellation techniques were used in the designs. The capacitive interpolation reduces the number of pre-amplifiers in the front (i.e., reducing the reference voltages).

The averaging technique decreases the non-linearity. The offset cancellation suppresses the offsets of pre-amplifiers and comparators. Four stages of pre-amplifiers, one stage of dynamic latched comparator and Grey code to prevent the sparkle code and metastability are used in the new flash ADC designs.

5.2 Future Work

Flash ADCs may be further improved by some design techniques. For example, a bandgap reference, which is widely used in data conversion systems, voltage regulators and memory systems, can be used in ADC designs to ensure very stable reference voltages.

The basic idea of BGR is to add a proportional to absolute temperature (PTAT) voltage to the emitter-base voltage (V_{BE}), so that the first-order temperature dependency of the pn junction is compensated by the PTAT voltage, and a nearly temperature independent output [35] is generated. The PTAT voltage is actually the thermal voltage (V_t) of the pn junction. V_{BE} or the forward voltage of a pn-juction diode exhibits a negative temperature coefficient while the difference between V_{BE}, say ΔV_{BE}, is directly proportional to the absolute temperature. Thus the reference voltage is expressed as [36]:

$$V_{ref} = V_{BE} + K \cdot V_t \tag{5.1}$$

If choosing a proper value of K, the reference is theoretically independent of temperature and external supply, illustrated below.

$$\frac{\partial V_{BE}}{\partial T} = \frac{V_{BE} - (4+m)V_T - E_g/q}{T} = -1.5mV/K$$

$$\frac{\partial \Delta V_{BE}}{\partial T} = \frac{\partial V_T \ln n}{\partial T} = \frac{k}{q}\ln n, \text{ where } \frac{\partial V_T}{\partial T} = +0.087mV/K. \quad (5.2)$$

$$\text{Therefore, } V_{ref} = V_{BE} + \frac{1.5}{0.087}V_T \approx V_{BE} + 17.2V_T \approx 1.25V$$

A simplified schematic of a factional BGR is illustrated in Figure 5.3 [36], where a fractional BGR is applied in sub 1-V applications.

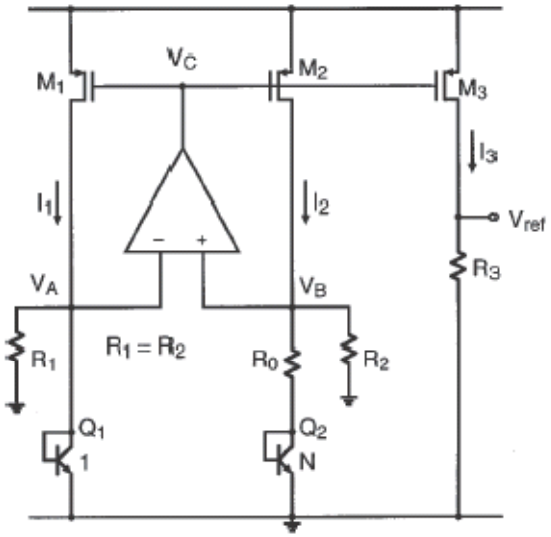


Figure 5.4 A Simplified Schematic of A BGR.

The PMOS size M_1, M_2 and M_3 are the same, R_1=R_2 and Q_2 is N times larger than Q_1 in size area. If M_1, M_2 and M_3 are all in the saturation mode, it is obvious I_1=I_2=I_3. The op-amp drives V_A and V_B equal, V_A=V_B. Thus I_{R1}=I_{R2}, and hence I_{Q1}=I_{Q2}. The VBE difference of the two BJT is,

$$\Delta V_Q = V_{Q1} - V_{Q2} = V_T \cdot \ln N$$
$$I_{R0} = \frac{V_{R0}}{R_0} = \frac{\Delta V_Q}{R_0} = \frac{V_T \cdot \ln N}{R_0} \tag{5.3}$$

The current across R_2 is,

$$I_{R2} = \frac{V_B}{R_2} = \frac{V_A}{R_2} = \frac{V_{Q1}}{R_2} \tag{5.4}$$

And $I_3 = I_2 = I_{R0} + I_{R2} = \frac{\Delta V_Q}{R_0} + \frac{V_{Q1}}{R_2}$. Therefore,

$$\begin{aligned} V_{ref} = I_3 \cdot R_3 &= \left(\frac{\Delta V_Q}{R_0} + \frac{V_{Q1}}{R_2} \right) \cdot R_3 \\ &= \frac{R_3}{R_2} \left(V_{Q1} + \frac{R_2}{R_0} \cdot \Delta V_Q \right) \\ &= \frac{R_3}{R_2} \cdot V_{ref\_conv} \end{aligned} \tag{5.5}$$

Thus if adjusting the resistance of R2 and R3, any portion of the conventional bandgap reference voltage could be obtained, even sub 1 – V voltage.

On the other hand, this new flash ADC design methodology can be further improved by taking into consideration of the higher order effects and building up more detailed and accurate model. Meanwhile, new capacitive interpolation technique with higher-than-two interpolation factor is under exploring for some scenarios which not very high speed (hundreds of MSps) ADC is needed but power consumption is priority instead of the speed. Because rough estimation manifests that if applying new capacitive

interpolation structure (interpolation factor higher than 2), fewer overall pre-amplifiers are used. Than total power dissipation may be reduced. Furthermore, a more general design technique may be developed for capacitive interpolation flash ADC not only with interpolation factor of 2 but also with interpolation factor higher than 2.

REFERENCE

[1] Pedro M. Figueiredo and Joao C. Vital, *Offset Reduction Techniques in High-Speed Analog-to-Digital Converters*, Springer, 2009.

[2] Mikael Gustavsson, J. Jacob Wikner and Nianxiong Nick Tan, *CMOS DATA CONVERTERS FOR COMMUNICATIONS*, Kluwer Academic Publishers, 2000.

[3] Jaehyun Lim, *Analog-to-Digital Converters*, Department of Computer Science and Engineering, the Pennsylvania State University, Mixed Signal Chip Design Lab.

[4] David William Cline, "*Noise, Speed and Power trade-offs in Pipelined Analog-to-Digital Converter*", University of California at Berkeley.

[5] Behzad Razavi, *Principles of Data Conversion System Design*, the Institute of Electrical and Electronics Engineers, Inc., New York, 1995.

[6] Yunchu Li, *Design of High Speed Folding and Interpolating Analog-to-Digital Converter*, Texas A&M University, May 2003.

[7] Boris Murmann, "*VLSI Data Conversion Circuit*", Department of Electrical Engineering, Stanford University, 2007.

[8] Haideh Khorramabadi, *Analog-Digital Interface Integrated Circuits*, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, 2007.

[9] Christoph Sandner, et al, “A 6-bit 1.2GSps Low-Power Flash ADC in 0.13μm Digital CMOS”, *IEEE Solid-State Circuits Conference*, Sept. 2004, pp.339 – 342.

[10] Pedro M. Figueiredo, *Member*, *IEEE*, and Joao C. Vital, *Member*, *IEEE*, “Kickback Noise Reduction Techniques for CMOS Latched Comparators”, *IEEE Transactions on Circuits and Systems-II: Express Briefs*, Vol. 53, No. 7, July 2006, pp541 – 545.

[11] Siqiang Fan, “High Speed ADC Design Techniques”, Electrical and Computer Engineering, Illinois Institute of Technology, 2007.

[12] J. Vandenbussche, et al., “Systematic design of a 200 MS/s 8-bit interpolating A/D converter”, *Proc. IEEE DATE*, 2002, pp. 357 – 361.

[13] Peter Xiao, et al, “A 4 b 8 GSample/s A/D converter in SiGe bipolar technology”, *IEEE Solid-State Circuits Conference*, February 1997, pp.124 – 125.

[14] A. Ismail and M. Elmasry, “Analysis of the flash ADC bandwidth-accuracy trade-off in deep-submicron CMOS technologies”, *IEEE Trans. Circuit and Systems-II*, 2008, pp. 1001 – 1005.

[15] Farhang Vessal, and C. Andre T. Salama, “An 8-Bit 2-Gsample/s Folding-Interpolating Analog-to-Digital Converter in SiGe Technology”, *IEEE Journal of Solid-State Circuits*, January 2004, pp.238 – 241.

[16] J. Lee, P. Roux, T. Link, Y. Baeyens and Y.-K. Chen, “10 Gsample/s optoelectronic A/D converter”, Electronics Letters, Nov. 2003, pp.1623 – 1624.

[17] Hiroshi Cimura, Akira Matsuzawa, Takashi Nakamura and Shigeki Sawada, “A 10-b 300-MHz Interpolated-Paralleled AD Converter”, *IEEE Journal of Solid-State Circuits*, April 1993, pp.438 – 446.

[18] Michael Choi, and Asad A. Abidi “A 6-b 1.3-Gsample/s A/D converter in 0.35-um CMOS”, *IEEE Journal of Solid-State Circuits*, Dec. 2001, pp. 1847 – 1858.

[19] Krishnaswamy Nagaraj, et al, “A dual-mode 700-Msamples/s 6-bit 200-Msamples/s 7-bit A/D converter in a 0.25-μm digital CMOS process”, *IEEE Journal of Solid-State Circuits*, December 2000, pp.1760 – 1768.

[20] Nathawad, Lalitkumar Y., Ryohei Urata, Bruce A. Wooley and David A. B. Miller. “A 40-GHz-bandwidth, 4-bit, time-interleaved A/D converter using photoconductive sampling”, *IEEE Journal of Solid-State Circuits*, Dec 2003, pp.2021 – 2030.

[21] Ken Poulton, et al, “A 20 GS/s 8 b ADC with a 1 MB memory in 0.18 um CMOS”, *IEEE International Solid-State Circuits Conference*, 2003, pp.318 – 496.

[22] Xicheng Jiang and Chang M.-C.F., “A 1-GHz signal bandwidth 6-bit CMOS ADC with power-efficient averaging”, *IEEE Journal of Solid State Circuits*, Feb. 2005, pp.532 – 535.

[23] Scholtens, Peter C. S., and Maarten Vertregt, “A 6-b 1.6-Gsample/s Flash ADC in 0.18-um CMOS Using Averaging Termination”, *IEEE Journal of Solid-State Circuits*, December 2002, pp.1599 – 1609.

[24] Johan van Valburg and Rudy J van de Plassche, “An 8-b 650-MHz Folding ADC”, *IEEE Journal of Solid-State Circuits*, December 1992, pp.1662 – 1666.

[25] Govert Geelen and Edward Paulus, “An 8b 600MSs 200mW CMOS Folding AD Converter Using an Amplifier Preset Technique”, *IEEE International Solid-State Circuits Conference*, February 2004, pp.254 – 526.

[26] Zheng-Yu Wang, Hui Pan, Chung-Ming Chang, Hai-Rong Yu and M. Frank Chang, “A 600 MSPS 8-bit Folding ADC in 0.18um CMOS”, *IEEE Symposium on VLSI Circuits*, June 2004, pp.424 – 427.

[27] Robert C. Taft, Chris A. Menkus, Maria Rosaria Tursi, Ols Hidri and Valerie Pons, “A 1.8-V 1.6-GSample/s 8-b Self-Calibrating Folding ADC With 7.26 ENOB at Nyquist Frequency”, *IEEE Journal of Solid-State Circuits*, December 2004, pp.2107 – 2115.

[28] Christoph Sandner, et al, “A 6-bit 1.2-GS/s low-power flash-ADC in 0.13-mu digital CMOS”, *IEEE Journal of Solid-State Circuits*, July 2005, pp.1499 – 1505.

[29] Hui Pan, “A 3.3-V 12-b 50-MS/s A/D Converter in 0.6-um CMOS with 80-dB SFDR”, PhD thesis. U of California, Los Angeles 1999.

[30] David F. Hoeschele, “Analog-to-Digital and Digital-to-Analog conversion Techniques”, 2nd edition, New York: John Wiley & Sons, Inc, 1994.

[31] Koen Uyttenhove and Michiel S. J. Steyaert, “A 1.8-V 6-Bit 1.3-GHz Flash ADC in 0.25-um CMOS”, *IEEE Journal of Solid-State Circuits*, July 2003, pp.1115 – 1122.

[32] William T. Colleran and A. A. Abidi, “A 10-b, 75-MHz two-stage pipelined bipolar A/D converter”, *IEEE Journal of Solid-State* Circuits, Dec. 1993, pp.1187 – 1199.

[33] Rudy J. van de Plassche, “Analog-To-Digital and Digital-To-Analog Converters”, New York: Springer, 1994.

[34] Dong-Young Chang, “Design techniques for a pipelined ADC without using a front-end sample-and-hold amplifier”, *IEEE Journal of Solid-State Circuits*, November 2004, pp.2123 – 2131.

[35] Xiaokang Guan, Albert Wang, et al, “A 3V 110uW 3.1ppm/ºC Curvature-Compensated CMOS Bandgap Reference”, *IEEE ISCAS*, Feb. 2006, pp. 2861 – 2864.

[36] Piero Malcovati, Franco Maloberti, *Fellow*, *IEEE*, Carlo Fiocchi, and Marcello Pruzzi, “Curvature-Compensated BiCMOS Bandgap with 1 – V Supply Voltage”, IEEE Journal of Solid-State Circuits, Vol. 36, No. 7, July 2001.

[37] Philip E. Allen and Douglas R. Holberg, *CMOS Analog Circuit Design*, Oxford University Press.

[38] K. Uyttenhove and M. Steyaert, “Speed-power-accuracy trade-off in high-speed ADC’s: what about nano-electronics?”, Proc. IEEE Custom Integrated Circuits Conf., 2001, pp. 341 – 344.

[39] W.-T. Lee, et al., “A new low power flash ADC using mutiple-selection method”, Proc. IEEE EDSSC, 2007, pp. 341 – 344.

[40] He Tang, Hui Zhao, Siqiang Fan, Xin Wang, Lin Lin, Qiang Fang, Jian Liu, Bin Zhao and Albert Wang, "Capacitive Interpolated Flash ADC Design Technique", *International SoC Design Conference*, Nov. 2010, pp.166 – 169.

[41] H. Tang, H. Zhao, S. Fan, X. Wang, L. Lin, Q. Fang, J. Liu, A. Wang and B. Zhao, "Design Technique for Interpolated Flash ADC", *IEEE ICSICT*, November 2010, Shanghai.

[42] He Tang, Hui Zhao, Siqiang Fan, Xin Wang, Lin Lin, Qiang Fang, Jian Liu and Albert Wang, "Design Matrix Analysis for Capacitive Interpolation Flash ADC", *Journal of Low Power Electronics*, accepted, 2010.

[43] He Tang, B. Qin, X. Wang, S. Fan, Q. Fang, L. Lin, J. Liu, J. He, H. Zhao, H. Tang and Albert Wang, "A Lower-Band UWB LNA with Integrated 7kV/15kV ESD Protection", *2nd IEEE Asia Symposium on Quality Electronic Design (ASQED),* pp.98-101, Penang, Malaysia, August 2010.

[44] He Tang, Chen Yang, Xin Wang, Siqiang Fan, Albert Wang and Tianling Ren, "On-Chip Magnetic Inductors for RFIC Applications", *18th International Conference on Composites/Nano Engineering (ICCE-18),* pp. 731-732, Anchorage, Alaska, July 2010.

[45] Nan Zhang, Xin Wang, He Tang, Albert Wang, Zhihua Wang and Baoyong Chi, "Low-Voltage and High-Speed FPGA I/O Cell Design in 90nm CMOS", *Proc. IEEE ASICON,* pp.533-536, 2009.