Pipeline ADC Design Methodology

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

Doctor of Philosophy

in

Electrical Engineering

by

Hui Zhao

June 2012

Dissertation Committee:

Prof. Albert Wang, Chairperson
Prof. Sheldon Tan
Prof. Qi Zhu
The Dissertation of Hui Zhao is approved:

________________________________________________
________________________________________________
________________________________________________

Committee Chairperson

University of California, Riverside
Acknowledgements

This research project would not have been possible without the support of many people. First and foremost, I offer my sincerest gratitude to my advisor, Dr Albert Wang, who has supported me throughout my thesis with his patience and knowledge. I attribute the level of my Ph.D. degree to his encouragement and effort and without him this thesis would not have been completed or written.

I wish to express my love and gratitude to my beloved families; for their understanding & endless love, through the duration of my studies.

Deepest gratitude is also due to my committee members: Prof. Dr Qi Zhu, Prof. Dr. Sheldon Tan and Prof. Albert Wang. I’d like to thank all my graduate friends, especially group members at LICS lab at UC Riverside, Xin Wang, He Tang, Qiang Fan, Zitao Shi, Lin Lin and Jian Liu. Without their knowledge and assistance, this study would not have been successful. Special thanks also to Zongyu Dong, Li Wang, and Rui Ma for teamwork and invaluable layout assistance.

I would also like to convey thanks to the OmniVision Technology Inc. for providing internship and test-chip tape-out opportunity.

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 grandmother.
Demand for high-performance analog-to-digital converter (ADC) integrated circuits (ICs) with optimal combined specifications of resolution, sampling rate and power consumption becomes dominant due to emerging applications in wireless communications, broad band transceivers, digital-intermediate frequency (IF) receivers and countless of digital devices. This research is dedicated to develop a pipeline ADC design methodology with minimum power dissipation, while keeping relatively high speed and high resolution.

Pipeline ADC is a mixed-signal system, which consists of sample and hold amplifier (SHA), sub-ADC, multiplying digital-to-analog Converter (MDAC) and bandgap voltage reference, comparator, switch-capacitor circuits and biasing circuits. This project set up a pipeline ADC design flow. It links all the specifications between the
system levels and circuit levels together. With this design flow, if the overall ADC specifications are given, such as resolution, sampling rate, voltage supply and input signal range, all the sub-block circuitry specifications are achieved.

This paper studies all the sub-block circuits of pipeline ADC first, and then come up with all the constraints and limitations for all the circuitry in term of speed and noises. Then a system level speed and power trade off consideration is explored in order to optimize the overall performance.

As verification of the proposed design methodology, a 10-bit 40MHz pipeline analog-to-digital converter prototype is developed in commercial TSMC 90nm CMOS technology: using op-amp sharing, dynamic biasing methods, it works in two modes: pipelined ADCs for high speed, cyclic ADC for low speed (only last stage runs, other stages are power off to save power). For pipeline mode, the total power consumption decrease as the sampling frequency drops.

Index terms: pipeline ADC, design methodology, CMOS, mixed-signal circuits
Contents

List of Figures x

List of Tables xiv

List of Symbols and Abbreviations xv

1 CHAPTER 1 INTRODUCTION 1

2 CHAPTER 2 OVERVIEW OF ADC 13

3 CHAPTER 3 PIPELINE ADC BLOCK STUDY 26

4 CHAPTER 4 PIPELINE ADC POWER OPTIMIZATION 73

5 CHAPTER 5 PIPELINE ADC DESIGN EXAMPLE 91

6 CHAPTER 6 CONCLUSIONS 110

REFERENCE 112
List of Figures

Figure 1.1 ADC in the Interface between Analog and Digital World.  2
Figure 1.2 ADC Applications.  4
Figure 1.3 Analog-to-Digital Converter Block Diagram.  5
Figure 1.4 ADC Static Errors.  9
Figure 1.5 ADC Dynamic Parameters.  12
Figure 2.1 A Typical Structure of A Flash ADC.  15
Figure 2.2 Two-Step Flash ADC.  18
Figure 2.3 First Order Sigma-Delta ADC.  20
Figure 2.4 A Pipelined ADC.  23
Figure 3.1 Pipeline ADC Transfer Curve.  27
Figure 3.2 (a) Stage Operation Modes (b) data latency in pipeline ADC  28
Figure 3.3 Detailed Pipeline ADC Architecture.  29
Figure 3.4 Pipeline ADC Design Matrix.  30
Figure 3.5 Pipeline ADC Design Flow.  32
Figure 3.6 (a) Comparator Symbol (b, c) Ideal/Practical Transform Function.  33
Figure 3.7 Amplifier-type Comparator.  34
Figure 3.8 (a) Typical Latch-type Comparator (b) Equivalent Small Signal Model.  35
Figure 3.9 Kickback Noise.  38
Figure 3.10 Static Latched Comparator.  39
Figure 3.11 Class-AB Latched Comparator.  39
Figure 3.12 Dynamic Latched Comparator.  40
Figure 3.13 Minimizing Kickback Noise.  43
Figure 3.14 (a) OP-AMP Notation, Ideal op-amp.  46
Figure 3.15 Two-stage Op-amps.  48
Figure 3.16 Telescopic Op-amp.                           49
Figure 3.17 Folded-cascode Op-amps.                        51
Figure 3.18(a, b) Gain Boosting Technology.                  .53
Figure 3.19 (a) SHA (b) Phase 1 (c) Phase 2.              56
Figure 3.20 (a) MDAC (b) Phase 1 (c) Phase 2.             62
Figure 3.21 Voltage Reference.                               64
Figure 3.22 Bandgap Voltage Reference.                      65
Figure 3.23 Curvature-compensated BGR.                      67
Figure 3.24 Reference Generator.                             69
Figure 4.1 Power F vs. R and n.                             75
Figure 4.2 optimized power vs. R.                            76
Figure 4.3 Capacitor Scaling Down.                           77
Figure 4.4 Normalized Power F vs. R and n.                  79
Figure 4.5 Optimized B vs. R.                               80
Figure 4.6 Optimized Power vs. R.                            82
Figure 4.7 Optimized Power with Both Capacitor and Resolution Scaling vs. R.         83
Figure 4.8 Pipeline ADC Structures.                       84
Figure 4.9 Power vs. Capacitor Scaling Based on Different Stage Partition.    90
Figure 5.1 Hybrid ADC.                                          94
Figure 5.2 Non-overlapping Clock Generator.                95
Figure 5.3 Comparator.                                            95
Figure 5.4 Dynamic Biasing.                                  96
Figure 5.5 (a) Folded-cascade Op-amp with Gain Boosting.    98
Figure 5.6 1st 2nd 3rd MDAC.                               99
Figure 5.7 4th MDAC.                                          100
Figure 5.8 Phase Scheme of Cyclic ADC.                       100
List of Tables

Table 2.1 Comparison of Different Types of ADCs. 25

Table 3.1 Performance Comparisons. 41

Table 3.2 Time Constant Comparisons. 42

Table 3.3 The Specifications of Two-stage Op-amp. 49

Table 3.4 The Specifications of Telescopic Op-amp. 50

Table 3.5 The Specifications of Folded-cascode Op-amp. 52

Table 3.6 The Specifications of Gain-boosting Op-amp. 54

Table 3.7 The Performance Comparisons of Op-amps. 55

Table 5.1 Hybrid Pipeline ADC Specification. 92

Table 5.2 Pins Definition of Pipeline ADC. 93
### List of Symbols and Abbreviations

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC</td>
<td>Analog-to-Digital Converter</td>
</tr>
<tr>
<td>BER</td>
<td>Bit Error Rate</td>
</tr>
<tr>
<td>BGR</td>
<td>Band-gap Reference</td>
</tr>
<tr>
<td>CMFB</td>
<td>Common-Mode Feedback</td>
</tr>
<tr>
<td>DAC</td>
<td>Digital-to-Analog Converter</td>
</tr>
<tr>
<td>DNL</td>
<td>Differential Non-Linearity</td>
</tr>
<tr>
<td>DR</td>
<td>Dynamic Range</td>
</tr>
<tr>
<td>DSP</td>
<td>Digital Signal Processing</td>
</tr>
<tr>
<td>ENOB</td>
<td>Effective Number Of Bits</td>
</tr>
<tr>
<td>FS</td>
<td>Full Scale</td>
</tr>
<tr>
<td>IF</td>
<td>Intermediate Frequency</td>
</tr>
<tr>
<td>INL</td>
<td>Integral Non-Linearity</td>
</tr>
<tr>
<td>LSB</td>
<td>Least Significant Bit</td>
</tr>
<tr>
<td>MDAC</td>
<td>Multiplying Digital-to-Analog Converter</td>
</tr>
<tr>
<td>MSB</td>
<td>Most Significant Bit</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>SC</td>
<td>Switched Capacitor</td>
</tr>
<tr>
<td>SFDR</td>
<td>Spurious-Free Dynamic Range</td>
</tr>
<tr>
<td>SNDR</td>
<td>Signal-to-Noise-and-Distortion Ratio</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>THD</td>
<td>Total Harmonic Distortion</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Motivation of ADC

In recent years, with the extensive implementation of digital computing and signal processing in communications, instrumentation, image processing and industrial control, analog to digital and digital to analog conversion applications in the field is expanding rapidly. Monolithic integration of technology, the novel circuit structure of integrated circuits, and advances in technology made the analog to digital and digital to analog conversion circuit design a great change, and continue to generate new areas of research and development.

At present, the electronic systems are in a trend of increasingly digitalization. Digital circuits and digital processing is almost everywhere. The main reason is: compared to the analog circuits, digital circuits with low noise sensitivity, strong anti-interference ability, high stability and wide adaptability, easy to design and automated testing, more extensive programmability features, but also because of advances in integrated circuit technology to continuously improve the performance of digital circuits. Large-scale integration (VLSI) technology allows each new generation of digital circuits to achieve higher speed, more features per chip, lower power consumption and cost. In addition, the progress of the improvement of the circuit structure and computer-aided
design (CAD) analysis and synthesis tools also promote the development of digital integrated circuits.

Although digital circuits advantages, and in many areas, it is gradually replacing analog circuit, but the physical environment of our existence cannot be separated from analog signal processing. First, the signals appearing in nature are almost always analog, such as temperature, pressure, time, speed, voltage, current, voice, luminous flux, etc. Secondly, the way of human perceives and keeps information is the analog mode.

In order to establish the interface of the digital processor and the analog world, data acquisition and reconstruction of the circuit is necessary. Therefore, Analog-to-Digital Convertor (ADC) and Digital-to-Analog Convertor (DAC) is very important and irreplaceable (Shown in Figure 1.1).

![Figure 1.1 ADC in the Interface between Analog and Digital World.](image)

ADC is a device that convert analog input signal into digital output signal. With the rapid development of digital computing and signal processing in electronic systems in the past few decades, it virtually transferred the world into a digital era. For a long time, as the bridge between the real analog world and the digital world, Analog-to-Digital Converters (ADCs) have been developed in different uni-directions to meet different
application requirements. Over-sampling sigma-delta ADCs, successive approximation
ADCs and algorithmic ADCs pushing the resolution up to 24bits, traded off with the
limited bandwidth, are widely employed in audio communication; Flash and folding and
interpolating ADCs, featuring highest speed at the cost of low resolution (no more than
10-bit) and large power dissipation and area, are adopted in disk drive channels and
magnetic data storage systems. Since 1990s, high-performance ADC ICs with optimized
combined specifications of resolution, sampling rate and power consumption (as opposed
to individual improvement) are in demand. For example, the IF frequencies in a typical
direct-IF (intermediate frequency) receiver vary from 50MHz to 200MHz, which requires
an ADC sampling rate up to 400MHz. Practically, the signal-to-noise ratio (SNR),
dynamic range and linearity requirements suggest that a better than 14 bits resolution is
necessary at affordable power dissipation. In addition, soft radio is proposed to digitize
signal at radio frequency (RF) domain and sort everything to the powerful digital signal
processing (DSP) IC chip, in order to bypass the obstacle originated from different
wireless standards. The critical building block in soft radio is the high sensitivity (that is
high SNR and high spurious free dynamic range (SFDR) ADC, with sampling frequency
being pushed further to several Giga-Hertz and relatively lower power consumption,
since the ADC must share the power budget with the other functional blocks in the same
transceiver system.
Market investigation results (shown in Figure 1.2) coming from commercial products of ADC manufactures (Analog Devices Inc.) indicate that ADC are widely used in many areas, such as communication, audio, computer, UWB, also the speed of ADC varies from less 100kHz to over GHz and the resolution varies from 6 to 22bits. To join with IC industry, this research focuses on the pipeline ADC whose resolution is 8~14bits and speed is 100k~200MHz in the mainstream CMOS process.
1.2 A/D Conversion Overview

1.2.1 ADC Definition

As the interface between the real world signals and the digital codes, ADCs play a major role in the data acquisition and digitizing process. ADCs convert continuous change of analog signal with amplitude and time into digital signals, which are discrete changes in the amplitude and time. Shown in Figure 1.3 is a block diagram of this process. Figure 1.3 shows the block diagram of an ADC, which consists of three building blocks – sampler, quantizer and coder. Input signals are converted to discrete time sampled signal by the sample circuit under the sampling clock control. Hold circuits play the role of the circuit to maintain the sampled value unchanged in the transformation process.

![Figure 1.3 Analog-to-Digital Converter Block Diagram.](image)

Sampler discretizes analog signal in time domain. The discrete-time continuous-amplitude signal is then mapped into a discrete level by quantizer, which implements a many-to-one transformer by approximating the signal to one of sub-divided quantization levels. Uniform quantization provides the minimum quantization error for a statistically equally distributed signal, while non-uniform quantization is a better choice for signals,
such as audio one, whose amplitude is not evenly distributed in the input range. For both the uniform and non-uniform quantization, a unique digital code is assigned to each quantization level. The analog-to-digital conversion process is fulfilled by providing a series of digital codes to represent the time-discrete and amplitude-discrete analog signal.

### 1.2.2 Main design direction of ADC performance

The first thing is high-speed. As the bandwidth of the signal source needs to be digitalized became wider and wider, so the audio, video, and RF ADC converters have to work at higher sampling rate.

Second is high-precision. Due to the growing dynamic range of the signal source needs to digitize, people have increasingly high quality requirements of data processing, so the accuracy of the converter continues to increase.

Third, it is easy for monolithic integration. At present, low-cost, low power consumption and high reliability of monolithic integration of the system on-chip system (SOC) has become a trend. Complete electronic systems are often mixed-signal systems, including digital circuits, analog circuits, analog to digital and digital to analog conversion circuit.

### 1.2.3 ADC Characterization

ADC performance is characterized by resolution, speed, power, INL, DNL and so on. It can be divided into two parts: static features and dynamic feature.
**Resolution**: ADC resolution is defined as the minimum input signal to make a change at digital output, usually expressed as the number of bits $n$ of the output binary code. The resolution shows the resolving ability of the input signal. Theoretically, for an $n$-bit ADC, the minimum input voltage that can be resolved is $V_{REF}/2^n$, equivalent to a quantitative unit of input voltage. Another term, called effective number of bits (ENOB), therefore, is introduced to measure the resolution for a pure sinusoidal signal accurately. It is defined as

$$ENOB = \frac{SNDR - 1.76}{6.02}$$  \hspace{1cm} (1.1)

where $SNDR$ represents signal-to-noise and distortion ratio.

- **Quantization error** (or quantization noise) is the difference between the original signal and the digitized signal. Hence, The magnitude of the quantization error at the sampling instant is between zero and half of one LSB. Quantization error is due to the finite resolution of the digital representation of the signal, and is an unavoidable imperfection in all types of ADCs.

- **Differential Non-linearity (DNL)**. DNL error is defined as the difference between an actual step width and the ideal value of 1LSB (see Figure 1.4). For an ideal ADC, in which the differential nonlinearity coincides with $DNL = 0$LSB, each analog step equals 1LSB ($1$LSB = $V_{FSR}/2^N$, where $V_{FSR}$ is the full-scale range and $N$ is the resolution of the ADC) and the transition values are spaced exactly 1LSB apart. A DNL error specification of less than or equal to 1LSB guarantees a
monotonic transfer function with no missing codes. An ADC's monotonicity is guaranteed when its digital output increases (or remains constant) with an increasing input signal, thereby avoiding sign changes in the slope of the transfer curve.

- **Integral Non-linearity (INL).** INL error is described as the deviation, in LSB or percent of full-scale range (FSR), of an actual transfer function from a straight line. The INL-error magnitude then depends directly on the position chosen for this straight line. At least two definitions are common: best straight-line INL and end-point INL. It is defined as the deviation of any single step from the ideal size and measured in LSBs. A less than ±1 LSB DNL error and less than ±1/2 LSB INL error specification guarantees no missing codes and a monotonic transfer function.

- **Offset Error.** Offset error identifies the horizontal difference between the actual transfer curve and the ideal one at the lowest transfer level. It is a figure-of-merit that measures the transfer curve matching at a single point and indicates the average error of the converter.

- **Gain Error.** Gain error, given as a percentage of the ideal input full-scale range, is reflected in the input/output transfer function as the slope deviation from the infinite resolution characteristic. Comparing with offset error, which can be measured at 1½ LSB above the most-negative end, gain error may be estimated at
1½ LSB below the most-positive end. Linear gain error doesn’t introduce distortion as does nonlinear gain error.

Figure 1.4 demonstrates the static errors in the same ADC transfer curve.

![Figure 1.4 ADC Static Errors.](image)

**Speed:** Analog-to-digital conversion speed is characterized by input bandwidth and sampling rate, which is defined as how many times the input analog signal is sampled per second and determined by the transient response parameters, such as settling time, slew rate and aperture jitter, of the internal blocks located in the analog signal path. The primary elements in the speed performance matrix are explained in this section.

Aperture jitter. When the input signal has large slewing rate, all the SNR of S/H circuit will decrease because of the special problem introduced by jitter issue from sampling clock. In order to get the relationship between the maximum allowable slewing rate and sampling rate, as well as resolution, let's consider that the voltage uncertainty of the input signal caused by clock jitter is less than 1LSB, so the jitter
impact could be ignored. For a full scaled input signal $\text{Vin} = A\sin 2\pi f t$, its maximum changing rate is $2\pi f maxA$, so:

$$2\pi f maxA \times \Delta t < 1\text{LSB}$$ (1.2)

$$2\pi f maxA \times \Delta t < \frac{2A}{2^n}$$ (1.3)

- **Total harmonic distortion (THD)**. THD is the ratio of the rms value of the fundamental signal to the mean value of the root-sum-square of its harmonics (generally, only the first 5 harmonics are significant). THD of an ADC is also generally specified with the input signal close to full-scale, although it can be specified at any level.

- **Signal-to-Noise Ratio (SNR)**. SNR (or sometimes called SNR-without-harmonics) is calculated from the FFT data the same as SINAD, except that the signal harmonics are excluded from the calculation, leaving only the noise terms. In practice, it is only necessary to exclude the first 5 harmonics, since they dominate. The SNR plot will degrade at high input frequencies, but generally not as rapidly as SINAD because of the exclusion of the harmonic terms. A few ADC data sheets somewhat loosely refer to SINAD as SNR, so you must be careful when interpreting these specifications and understand exactly what the manufacturer means. In a well-designed and matching converter, SNR presents an upper limit to both the static and dynamic performance and is characterized by a single tone or a multi tone measurement with a full scale sinusoidal input as
\[ SNR_{dB} = 6.02N + 4.77 - 20\log \frac{\text{Signal \_ peak \_ amplitude}}{\text{Signal \_ rms \_ value}} \]  

(1.4)

Where \( N \) is the resolution of ADCs. If the signal level is reduced, the value of SNR decreases and the ENOB decreases. It is necessary to add a correction factor for calculating ENOB at reduced signal amplitudes as shown in Equation (1.4)

- **Signal-to-Noise and Distortion Ratio (SNDR).** SNDR is the ratio of the rms value of the fundamental signal to the mean value of the root-sum-square of its harmonics plus all noise components (excluding dc). The bandwidth over which the noise is measured must be specified. In the case of an FFT, the bandwidth is dc to fs/2. (If the bandwidth of the measurement is dc to fs/2 (the Nyquist bandwidth), THD + N is equal to SINAD—see below). Be warned, however, that in audio applications the measurement bandwidth may not necessarily be the Nyquist bandwidth. The SINAD plot shows that the ac performance of the ADC degrades due to high-frequency distortion and is usually plotted for frequencies well above the Nyquist frequency so that performance in undersampling applications can be evaluated. SINAD plots such as these are very useful in evaluating the dynamic performance of ADCs.

- **Spurious Free Dynamic Range (SFDR).** Spurious free dynamic range is the ratio of the rms value of the signal to the rms value of the worst spurious signal regardless of where it falls in the frequency spectrum. The worst spur may or may not be a harmonic of the original signal. SFDR is an important specification in communications systems because it represents the smallest value of signal that
can be distinguished from a large interfering signal (blocker). SFDR can be specified with respect to full-scale (dBFS) or with respect to the actual signal amplitude (dBc).

Figure 1.5 illustrates the relationship among these dynamic parameters.
CHAPTER 2

ADC was first boosted in 1930s because of widely used PCM (Pulse Code Modulation) technology in telecommunication area. It changed the analog coding and decoding technology into digital signal processing technology. Until 1950s, with the emergence of high-speed digital computer and aircraft/missile data processing system, ADC has gained a further development.

After Microprocessors were well-developed in 1970s, it triggered a technological revolution in the field of digital signal processing and computer. In 1971, the first Monolithic ADC was designed as analog/digital interface circuitry. In the past 20 years, due to deep sub-micron integrated technology becomes more popular; it promotes a new area of analog integrated circuits: mixed signal integrated circuit. This chapter will briefly introduce all typical ADC, such as flash ADC, two-step ADC, pipeline ADC and so on.

OVERVIEW OF ADC

2.1 AD Classification

According to relationship between sampling rate and signal frequency, ADC can be divided into three types: the Nyquist ADC, subsampling ADC, and oversampling ADC.
Nyquist ADC: in order to ensure accurate reconstruction of the original value theoretically, ADC must comply with the sampling theorem, the sampling frequency is greater than or equal to twice the highest frequency of the input signal. Since anti-aliasing filter cannot be an ideal low-pass characteristics, there must be a transition band, so the sampling frequency slightly higher than twice the bandwidth of the analog signal.

Sub-sampling ADC: input signal only occupies a small part of the band pass frequency, it is possible to make the sampling frequency smaller than the highest signal frequency, but to ensure that the sampled spectrum do not overlap.

Oversampling ADC, the sampling frequency is much higher than the Nyquist frequency. The following part of the digital filter circuit is used to remove noise outside the signal bandwidth. Oversampling technique can reduce the quantization noise level, in order to achieve high-precision.

According to performance, ADC can be divided into high-speed AD converter, and high accuracy AD converter.

Based on architecture differences, ADC can be divided into the serial structure (pipeline), parallel structure (time-interleaving), and serial and parallel AD converter.

2.2 Flash AD Converters

Flash converters are extremely fast compared to many other types of ADCs. It is also quite simple and, apart from the analog comparators, only requires logic for the final
conversion to binary. It can achieve extremely high speed (over 1Gsamples/second) with low resolution.

The typical structure of a 2bit flash ADC is shown in Figure 2.1. A Flash converter requires a huge number of comparators compared to other ADCs, especially as the precision increases.

![2-Bit Flash Analog to Digital Converter](image)

**Figure 2.1 A Typical Structure of A Flash ADC.**

An N-bit flash ADC needs a resistor ladder composing $2^N$ equal resistors. These resistor ladders generate all possible $2^N-1$ voltage levels, then followed by $2^N-1$ comparators and digital encoders. Comparators are usually composed of a pre-amplifier used for sampling and amplifying the input signal, and a latch used for making final decision.
Each comparator samples the input signal and compares this signal to 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. Finally, 1-of-n codes are converted to binary outputs.

This ADC requires a lot of the comparator, and thus the hardware cost of the power, area and other large input capacitance. Its resolution is limited by the accuracy of the reference voltage and the offset voltage of the comparator. For CMOS technology, it also need to trade-off between the small offset voltage and high conversion speed. Therefore, the structure is more commonly used in the occasion of the 6 ~ 8 precision. Flash ADC is the most classic high-speed AD converter, and other structures in the high-speed AD converter are evolved on this basis.

The flash AD converter key weakness is that the hardware cost has grown exponentially with the resolution. The number of comparators is \(2^n-1\). When the resolution is 8 or more, it requires very large power and area. Also, because the heavy used comparators and lack of front end sample and hold amplifier, it will bring some special non-ideal characteristics, such as the variations of the ladder reference voltage, the interference immunity of signal glitches variation, the sensitivity of sampling point of different input signal slew rate. These issues will impact the conversion accuracy.
When the difference between two input signals of the comparator is relative small, it will take longer period of time to go through in order to decide a stable logic output. This condition is called comparator metastability. If the instantaneous value of the input signal value of flash AD converter is very close to one of comparators reference voltage, then the comparator output will be a long time uncertain, may bring the wrong digital output.

We can take some approaches to reduce the probability of occurrence of metastable error. For example: adding the number of latch behind each comparator, so as to allow a longer time to rebuild the thermometer code, which is a simple solution, but the power consumption is higher. Thermometer code to binary decoding using pipelining method is another way; this method allows the output of the comparator to experience more reconstruction time, but each comparator output level as the input of a logic gate. In addition, using Gray code between thermometer code and binary code can also effectively inhibit the comparator metastability error. In Gray coding circuit, the input signal does not appear in more than one input gate, allowing to increase the time of reconstruction with pipelining.

2.3 Two-Step Flash ADC

Because of the exponential growth of power dissipation, area, and input capacitance of a flash convert makes it impractical for resolution above 8 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 and an inter-stage gain amplifier between the subtractor and the fine flash ADC are necessary. The block diagram in Figure 2.2.1 illustrates the structure of a two-step flash ADC.

![Figure 2.2 Two-Step Flash ADC.](image)

In this type of ADC, 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 resolution of the first coarse flash ADC is $m$ bit, and that of second fine flash ADC is $n$ bit. So the total digital output result is $m+n$. If digital correction is used, one of the flash ADC need residual bit, so the total resolution is larger $m$ (or $n$) bit but smaller
than m+n bit. In flash ADC, comparators track the input signal for half of the sampling period, while make conversion for the other half of sampling period. However, for two-step ADC, because of extra DA conversion, subtraction, and second flash ADC work at same time, the speed of two-step ADC is limited.

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. Although the speed the two-step ADC is lower than flash ones. But the number of comparators is much smaller than that of flash ADC with same resolution: from $2^{n+m}-1$ to $2^n+2^m-2$. So the power and area are decreased.

### 2.4 Oversampling AD Converters

An oversampling converter uses a noise-shaping modulator to reduce the in-band quantization noise to achieve a high degree of resolution. It can shape the quantization noise and push the majority of the inband noise to higher frequencies. It modulates the analog input signal to a simple digital code, normally a one-bit serial stream using a sampling rate much higher than the Nyquist rate.

To understand noise shaping, consider the block diagram of a sigma-delta modulator of the first order (Figure 2.3). It includes a difference amplifier, an integrator, and a comparator with feedback loop that contains a 1-bit DAC. (This DAC is simply a switch that connects the negative input of the difference amplifier to a positive or a negative reference voltage.) The purpose of the feedback DAC is to maintain the average output of the integrator near the comparator's reference level.
The density of "ones" at the modulator output is proportional to the input signal. For an increasing input the comparator generates a greater number of "ones," and vice versa for a decreasing input. By summing the error voltage, the integrator acts as a low-pass filter to the input signal and a high-pass filter to the quantization noise. Thus, most of the quantization noise is pushed into higher frequencies. Oversampling has changed not the total noise power, but its distribution.

If we apply a digital filter to the noise-shaped delta-sigma modulator, it removes more noise than does simple oversampling. This type of modulator (first-order) provides a 9dB improvement in SNR for every doubling of the sampling rate. For higher orders of quantization, we can achieve noise shaping by including more than one stage of integration and summing in the sigma-delta modulator.

![Diagram of first order sigma-delta ADC](image)

**Figure 2.3 First Order Sigma-Delta ADC.**

The oversampling ratio, called M, is a ratio of the clock frequency to the Nyquist frequency of the input signal. This oversampling ratio can vary from 8 to 256. The resolution of the oversampled converter is proportional to the oversampled ratio. The bandwidth of the input signal is inversely proportional to the oversampled ratio.
It is very compatible with VLSI technology because most of the converter is digital. High resolution. Single-bit quantizers use a one-bit DAC which has no INL or DNL errors. Provide an excellent means of trading precision for speed (16-18 bits at 50 ksp to 8-10 bits at sampling rates of 5-10 Msps).

Oversampled ADCs allow signal bandwidth to be efficiently traded for resolution. Noise shaping oversampled ADCs preserve the signal spectrum and shape the noise quantization spectrum. The modulator shapes the noise quantization spectrum with a high pass filter. This high-pass characteristic reduces the noise at low frequencies which is the key to extending the dynamic range within the bandwidth of the converter. The quantizer can be single or multiple bits. Single bit quantizers do not require linear DACs because a 1 bit DAC cannot be nonlinear. Multiple bit quantizers require ultra linear DACs. Modulators consist of combined integrators with the goal of high-pass shaping of the noise spectrum and cancellation of all quantizer noise but the last quantizer.

2.5 Pipelined ADC

The pipelined analog-to-digital converter (ADC) has become the most popular ADC architecture for sampling rates from a few megasamples per second (Msps) up to 100 Msps. Compared to the two-step flash ADC which has just two stages, pipeline ADCs have multiple cascades stages.
Resolutions range from eight bits at the faster sample rates up to 16 bits at the lower rates. These resolutions and sampling rates cover a wide range of applications, including CCD imaging, ultrasonic medical imaging, digital receivers, base stations, digital video (for example, HDTV), xDSL, cable modems, and fast Ethernet.

Applications with lower sampling rates are still the domain of the successive approximation register (SAR) and integrating architectures, and more recently, oversampling/sigma-delta ADCs. The highest sampling rates (a few hundred Msp or higher) are still obtained using flash ADCs. Nonetheless, pipelined ADCs of various forms have improved greatly in speed, resolution, dynamic performance, and low power in recent years.

Each stage of the pipeline 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 pipeline ADC is illustrated in Figure 2.4.
Figure 2.4 A Pipelined ADC.

In this schematic, the analog input, IN, is first sampled and held steady by a sample-and-hold (S&H), while the flash ADC in stage one quantizes it to $n_1$ bits. The $n_1$ bit output is then fed to a $n_1$-bit DAC, and the analog output is subtracted from the input. This "residue" is then gained up by a factor of $G_1$ and fed to the next stage (Stage 2). This gained-up residue continues through the pipeline, providing $n_i$ bits per stage until it reaches the $n_k$-bit flash ADC, which resolves the last $B_k$ bits. Because the bits from each stage are determined at different points in time, all the bits corresponding to the same sample are time-aligned with shift registers before being fed to the digital-error-
correction logic. Note when a stage finishes processing a sample, determining the bits, and passing the residue to the next stage, it can then start processing the next sample received from the sample-and-hold embedded within each stage. This pipelining action is the reason for the high throughput.

The concurrency of the pipeline 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 pipeline 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.

2.6 Summary of ADCs

Table 2.1 shows the differences of different ADCs. Flash ADC has the fast speed while lowest resolution. Oversampling (Sigma-delta) ADC can achieve over 20 bits resolution, but the sampling rate is lowest. Two-step flash ADC and pipeline ADC sit in the middle. However, pipeline ADC can get higher resolution but lower speech compared to two-step ADC.

So based on different application and requirement, ADC structure might not be the same. Flash ADC focus on radio, UWB, WiFi applications because of its high speed features. Two-step and pipeline ADC will be used for communication, video, image
sensor, and baseband systems. While oversampling ADC will be used for audio, communication, thermal sensing and precision test systems.

<table>
<thead>
<tr>
<th>ADC</th>
<th>Resolution</th>
<th>SNR/dB</th>
<th>Fs/Hz</th>
<th>Power</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flash</td>
<td>4~6</td>
<td>30~46</td>
<td>&gt;1G</td>
<td>higher</td>
<td>Radio, UWB, WiFi, high speed system</td>
</tr>
<tr>
<td>Two-step</td>
<td>8~10</td>
<td>44~60</td>
<td>500M</td>
<td>High</td>
<td>Communication, video</td>
</tr>
<tr>
<td>Pipeline</td>
<td>8~16</td>
<td>48~80</td>
<td>1M~200M</td>
<td>low</td>
<td>Communication, video, baseband system</td>
</tr>
<tr>
<td>Sigma-Delta</td>
<td>16~24</td>
<td>90~130</td>
<td>10M</td>
<td>lower</td>
<td>Audio, communications, precision test system</td>
</tr>
</tbody>
</table>

Table 2.1 Comparison of Different Types of ADCs.
CHAPTER 3

PIPELINE ADC BLOCK STUDY

Pipeline ADC uses two or more steps of subranging. First, a coarse conversion is done. In a second step, the difference to the input signal is determined with a digital to analog converter (DAC). This difference is then converted finer, and the results are combined in a last step. This can be considered a refinement of the successive-approximation ADC wherein the feedback reference signal consists of the interim conversion of a whole range of bits (for example, four bits) rather than just the next-most-significant bit. By combining the merits of the successive approximation and flash ADCs this type is fast, has a high resolution, and only requires a small die size.

3.1 Basic Concept of Pipeline ADC

The principle of sub-ranging ADC can be pushed to the limit of having only one bit per stage. At this point, each flash ADC is nothing more than a simple comparator; also, the data is transferred in a pipeline fashion: when the data is sent to the second stage, another sampled data is fed to the first stage; the result is a latency delay equal to the number of stages. Since the pipeline ADC is pipelining the subranging structure, and the binary search in the sub-ADCs runs just as the mathematic division, the first stage decides the MSBs and the last stage sets the LSBs. The MSBs divide the full reference
range, while LSBs divide the sub reference range. The relationship between MSBs and LSBs is revealed in Figure 3.1.

![Figure 3.1 Pipeline ADC Transfer Curve.](image)

A Pipeline ADC consists of a cascade of stages, each of which contains a low resolution ADC, DAC and amplifier, which successively convert the analog input into its digital representation, while processing the data in a pipe-lined manner.

Pipeline ADCs are commonly used for power-efficient high-speed conversion of wide bandwidth input signals (e.g. 10 to 100 MHz). The ADC sampling frequency is usually the Nyquist frequency or lower using small OSRs (e.g. 2 or 4) and the ADC output code resolution is typically between 8 and 14-bit.

As a well-organized data processing system, the operation of the pipeline ADCs is under stringent timing control. Each stage performs data conversion in sampling and holding modes serially. The sampling and holding modes interleave between two adjacent stages and the digital output is valid after some clock cycles, called latency time, decided by the number of stages. This process is illustrated in Figure 3.2(a). Because
each sample must propagate through the entire pipeline before all its associated bits are available for combining in the digital-error-correction logic, data latency is associated with pipelined ADCs. In the example in Figure 3.2(b), this latency is about seven cycles.

![Diagram of ADC stages with sampling and holding modes](image)

**Figure 3.2** (a) Stage Operation Modes in Pipeline ADC (b) data latency in pipeline ADC.
This timing scheme of the pipeline ADC is built up by the sample and hold circuit in each stage. During sampling mode, the switch controlled by the sampling clock is connected to the residue generated from the preceding stage and the signal is sampled on the sampling capacitor. When the hold clock comes, the switch is turned off and the signal stored on the sampling capacitor on one hand is converted to the thermometer codes by the sub-flash ADC, on the other hand, subtracts the estimated analog signal that is re-constructed by the D/A converter, to create the new residue as the input signal of the following stage. The thermometer codes from each stage are encoded to the binary ones and latched and added together to form the final m-bit digital output (where m is the resolution of the pipeline ADC). The detailed pipeline architecture is shown in Figure 3.3.

Pipeline ADC design includes system level design and block level design. For system level consideration, it consists of architecture, stages partition, power optimization, specification of sub-block circuits. Designer should consider what kind of
architecture the pipeline ADC should be: op-amp sharing, S/H circuits, dithering or not, how many bit one stage of MDAC should handle, how to minimize power and what the specification of sub-block should be, like op-amp DC gain and close-loop gain bandwidth.

So, it’s a complex design trade-off matrix among resolution, thermal noise, power, input range, voltage supply, op-amp types, nonlinearity, and so on. It is shown on Figure 3.4. This work comes up with an optimized methodology of pipeline ADC design trade-off and finds all the connection among these important features. For example, as resolution increase, the requirement of thermal noise, which is related to KT/C, also become tight (SNR also increase). So as to minimize thermal noise, the sample capacitance in MDAC will increase; however, it will burn more power. If the input range goes up, the requirement of thermal noise will drop, but the selection of op-amp becomes to be one dominant issue.

![Figure 3.4 Pipeline ADC Design Matrix.](image-url)
So, the methodology study of pipeline ADC design flow shows as following figures 3.5.

1.) Before designing pipeline ADC, first thing to know all the specification of ADC, like resolution, input signal range, maximum sampling rate, voltage supply and so on.

2.) Then, decide what kind of ADC architecture will be used, number of stages, sampling capacitor value.

3.) Define the block level circuits specification based on the first two steps. Like op-amp types, DC gain, closed-loop gain bandwidth, switch selections and so on.

4.) Design and simulation the block level circuits to beat the specification.

5.) Design and simulation the top level ADC.

6.) Run post simulation and corner, Monte Carlo simulation.

7.) If everything passes, the layout could be tape-out.

All the block level design is important, so we first study the block level circuits and will discuss the system level power optimization in the following chapter.
Figure 3.5 Pipeline ADC Design Flow.
3.2 Comparator

A comparator is a device which compares two voltages or currents and switches its output to indicate which is larger. The most widely used ADCs process voltage signal, so here we only discuss voltage comparator. Its symbol is in figure 3.6 (a) and its ideal transform function is figure 3.6(b).

\[ V_{\text{out}} = \begin{cases} V_{S+} & V_1 > V_2 \\ V_{S-} & V_1 < V_2 \end{cases} \]

(a)

Figure 3.6 (a) Comparator Symbol (b) Ideal Transform Function (c) Practical Transform function.

The important Specifications of voltage comparator are gain $A_V$, offset $V_{OS}$, speed (step response), kickback noise and power consumption. The gain of ideal comparator is infinite, however the practical comparator gain is:
\[ A_V = \frac{V_{OH} - V_{OL}}{V_{IH} - V_{IH}} \]  

\( V_{OS} \) is the input voltage necessary to make the output equal to half of \((V_{OH}+V_{OL})\) when \(V_P=V_N\). Comparator speed, kickback noise and power issue will be discussed in details in the following sections. There are mainly two type of voltage comparator: amplifier-type comparator and Latch-type comparator which is also called regenerative comparator.

### 3.2.1 Amplifier-type comparator

The natural choice to design a comparator is to use an amplifier. A small voltage at the input is then amplified to a value large enough to be detected by the following digital logic circuits.

![Amplifier-type Comparator Circuit](image)

**Figure 3.7 Amplifier-type Comparator.**

For pipeline/flash ADC design, we seldom use amplifier-type comparator. The reasons are:
1. It will always consume static power which means it is costly. The power of comparator in Figure 3.7 is $P=I_{M5} \cdot V_{DD}+I_{M7} \cdot V_{DD}$.

2. The speed of amplifier-type comparator is slow, compared to latch-type comparator, there is no positive feedback in amplifier-type one, so the time response is slow. As a result it’s difficult to achieve the performance needed for a 12 bit 50 MS/s AD converter [1].

### 3.2.2 Latch-type comparator

Figure 3.8 shows a typical latch-type comparator [9]. The most important feature of this type comparator is that the output of one amplifier is connected to the input of the other amplifier. These cross-couple connection makes a positive feedback in that loop, which results in faster time response compared with amplifier-type one.

![Typical Latch-type Comparator](image1)

![Equivalent Small Signal Model](image2)

Figure 3.8 (a) Typical Latch-type Comparator (b) Equivalent Small Signal Model.
The input differential voltage is applied to a source-coupled pair, which produces signal currents that are loaded by the resistance $R_{SW}/R_{RGN}$. $R_{SW}$ is the drain-source resistance of $M_5$ (drain-source resistances of $M_1$-$M_4$ are neglected). $R_{RGN}$ is the resistance due to the regeneration of $M_3$-$M_4$ which has been shown to be equal to $-2/g_{m3}$. $C_P$ represents parasitic capacitance on node $V_{d3}$ and $V_{d4}$. The comparator works in two non-overlapped phases.

Phase one is reset mode: when the signal $V_{SW}$ is high, the reset mode requires that the value of $R_{SW}/R_{RGN}$ be positive, which will be the case if $M_5$ is in triode and $R_{SW} < 2/g_{m3}$. In this mode, the stage behaves as a stable, low-gain differential amplifier with $(v_{d4} - v_{d3})$ being the differential output voltage. Phase two is regeneration mode: When $V_{SW}$ goes low, $R_{SW}$ becomes very large and hence $R_{SW}/R_{RGN}$ becomes negative. Under this condition, the stage is unstable and the magnitude of the output voltage will increase exponentially with time due to the negative time constants at the nodes $v_{d3}$ and $v_{d4}$. Hence, this is called the regeneration mode. In operation, when the clock transitions from high to low any existing output voltage will regenerate until either $v_{d3}$ or $v_{d4}$ goes low and the other goes high.

Considering symmetry, the input current source $i_{in}$ represents the drain current of $M_1$, the capacitor $C_p$ represents the parasitic capacitance at $v_{d3}$, and the resistor $R$ represents one-half of $R_{SW}/R_{RGN}$.

- Reset mode:
  - $R = R_1 = (R_{SW}/R_{RGN})/2 > 0$
- Time constant

\[ \tau_1 = C_p R_1 = \frac{R_{SW} || R_{RGN}}{2} C_p > 0 \]  

(3.2)

- Output voltage

\[ \Delta V_0 = A\Delta V_{in} (1 - e^{\frac{t}{\tau_1}}) \]  

(3.3)

- regeneration mode:

\[ R = R_2 = R_{RGN}/2 = -1/g_{m3} < 0 \]

- Time constant

\[ \tau_2 = C_p R_2 = -\frac{1}{g_{m3}} C_p < 0 \]  

(3.4)

- Output voltage

\[ \Delta V = \Delta V_0 e^{\omega_d t} = \Delta V_0 e^{\frac{t}{\tau_2}} \]  

(3.5)

The final output voltage of these two phases is:

\[ \Delta V_{final} = \Delta V_0 e^{\frac{t}{\tau_2}} = A\Delta V_{in} (1 - e^{\frac{t}{\tau_1}}) e^{\frac{t}{\tau_2}} \]  

(3.6)

A smaller regeneration time constant \( \tau_2 \) will create a full-scale output sooner. But it will reduce the sensitivity of the comparator to inputs after the start of regeneration. The reset time constant \( \tau_1 \): A longer reset time constant will increase the sensitivity of the comparator to inputs before the start of regeneration.
There are mainly 3 types of latch-type comparator [10]: Static latched comparator, Class-AB latched comparator and Dynamic latched comparator. Before discuss these comparators, kickback noise is first introduced in figure 3.9: The large voltage variations on the regeneration nodes are coupled, through the parasitic capacitances of the transistors, to the input of the comparator. Since the circuit preceding it does not have zero output impedance, the input voltage is disturbed, which may degrade the accuracy of the converter. This disturbance is usually called kickback noise.

![Figure 3.9 Kickback Noise.](image)

### 3.2.3 Static latched comparator

The regeneration is done by two cross-coupled inverters (M\textsubscript{3a}/M\textsubscript{4a} and M\textsubscript{3b}/M\textsubscript{4b}). The power efficiency is poor, since the consumption is purely static. But it has low kickback noise because of slow regeneration process.
3.2.4 Class-AB latched comparator

Figure 3.11 Class-AB Latched Comparator.
The regeneration is done by two cross-coupled CMOS inverters. Their current increases momentarily, during the regeneration process, to charge the output nodes faster. The drains of the input differential pair are directly connected to the regeneration nodes. The circuit reacts quicker to input variations, because there is only one pole. However, this increases the kickback noise; it still has supply current in the reset phase and after the regeneration finishes.

3.2.5 Dynamic latched comparator

![Dynamic Latched Comparator Diagram]

Figure 3.12 Dynamic Latched Comparator.

In dynamic latched comparators [10], the current only flows during the regeneration. After regeneration is completed, one output node is at $V_{DD}$; the other output and both drains of the differential pair have a 0-V potential. There is no supply current. This comparator has two kickback noise sources: voltage variations and the variation of the operating region of the differential pair transistors. In the reset mode: all the
transistors are cut-off (The gate, source and drain of M₂a and M₂b are all at Vᵩᵩ). In the regeneration mode: at the very beginning, M₁a and M₁b are in saturation; M₂a and M₂b are cut-off, so no current runs through Vᵩᵩ to ground. The current discharges parasitic capacitors Cₚ₁ at the M₁’s drain. Until one of the transistor M₂ is turn on. When one of M₂ turns on, the voltage of its drain will decrease first, leading the opposite PMOS turns on. Then the regeneration begins to work.

Table 3.1 shows all the comparisons among those four types of comparators. And table 3.2 compares the time response of these comparators which indicate the speed performance.

<table>
<thead>
<tr>
<th>Performance</th>
<th>Speed</th>
<th>Positive feedback</th>
<th>Power</th>
<th>Kickback noise</th>
<th>Area transistors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplifier</td>
<td>Slowest</td>
<td>None</td>
<td>Most</td>
<td>None</td>
<td>7</td>
</tr>
<tr>
<td>Static latch</td>
<td>Slow</td>
<td>Y</td>
<td>More</td>
<td>Less</td>
<td>11</td>
</tr>
<tr>
<td>Class-AB latch</td>
<td>Faster</td>
<td>Y</td>
<td>Less</td>
<td>More</td>
<td>9</td>
</tr>
<tr>
<td>Dynamic latch</td>
<td>Fastest</td>
<td>Y</td>
<td>Least</td>
<td>Most</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 3.1 Performance Comparisons.
Table 3.2 Time Constant Comparisons.

<table>
<thead>
<tr>
<th>Time constant</th>
<th>$\tau_1$</th>
<th>$\tau_2$</th>
<th>Poles in pre-amp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static latch</td>
<td>$\frac{1}{g_{m_2}}C_{p_1} + \frac{1}{g_{m_3}}C_{p_2}$</td>
<td>$\frac{1}{g_{m_3}}C_p$</td>
<td>2</td>
</tr>
<tr>
<td>Class-AB latch</td>
<td>$\frac{R_{eq}}{2}C_p$</td>
<td>$\frac{1}{g_{m_3}+g_{m_2}}C_p$</td>
<td>1</td>
</tr>
<tr>
<td>Dynamic latch</td>
<td>---</td>
<td>$-(\frac{1}{g_{m_3}+g_{m_2}})C_{p_2}$</td>
<td>1</td>
</tr>
</tbody>
</table>

From these tables, dynamic latch comparator is the fastest one and consumes least power. It doesn’t even need reset time because of no RC time constant in that phase, but it has the largest kickback noise. Amplifier-type comparator is the most costly since it has a static current run through $V_{DD}$ to ground all the time, however, it doesn’t has kickback noise for no positive feedback loop exists.

### 3.2.6 Non-ideal problems in comparator

Offset and kickback noise interference are the mainly problems in comparators. In order to get offset as small as possible, one way is minimization; and anther is compensation or cancellation, that mechanism is much similar to offset compensation in OP-AMP, which will discussed in that section.

In order to minimize offset, pre-amplifier which is prior to comparator is introduced. The total offset is:

$$V_{os,in_2} = V_{os,in,amp} + \frac{V_{os,in,latch}}{A}$$  \hspace{1cm} (3.7)
The offset of latch is minimized by the factor of A (A is the gain of pre-amplifier). As we all due to the well- symmetric topology of pre-amplifier, the pre-amplifier offset is rather smaller than latch one. We can also increase the area of pre-amplifier to get a less mismatching.

Isolation can minimize the voltage variations on the drains of the differential pair [10]. Those nodes are isolated from the regeneration nodes using switches, which open during the regeneration phase. An alternative path for the current of the differential pair must be provided, in order to keep the drain voltages near the values found in the reset phase. (Figure 3.13)

![Figure 3.13 Minimizing Kickback Noise.](image-url)
3.2.7 Specification abstraction

Offset requirement: the offset of comparator is required less than 1LSB in order to achieve no missing code. $V_{os} < V_{ref}/2^N$

Speed requirement:

1. Preamplifier:

Any signal which is larger than 1LSB, should be amplified during reset mode.

(a) For small differential input signal, which is a little bit larger than 1LSB: if the pre-amplifier has only one pole:

$$t_{setting} = \tau_1 \ln \frac{1}{1 - \frac{\Delta V_0}{A}} < \frac{1}{2f_{sample}}$$

(3.8)

(b) For large differential input signal,

$$t_{setting} = \Delta V \frac{\Delta V \times C_p}{SR} \frac{1}{I} < \frac{1}{2f_{sample}}$$

(3.9)

Sampling rate:

$$f_{sample} < \min \left\{ \frac{I}{2\Delta V \times C_p}, \frac{1}{2\tau_1 \ln \frac{1}{1 - \frac{\Delta V_0}{A}}} \right\}$$

(3.10)
2. Latch

\[ \Delta V = \Delta V_0 e^{o_s t} = \Delta V_0 e^{\tau_2} > V_{OH} - V_{OL} \Rightarrow t_{setting} > \tau_2 \ln \left( \frac{V_{OH} - V_{OL}}{\Delta V_0} \right) \]

\[ \Rightarrow f_{sample} < \frac{1}{2t_{setting}} = \frac{1}{2\tau_2 \ln \left( \frac{V_{OH} - V_{OL}}{\Delta V_0} \right)} \]  

(3.11)

As a result, the final specifications of comparator are:

\[ f_{sample} < \min \left\{ \frac{I}{2 \Delta V \times C_p}, \frac{1}{2\tau_1 \ln \frac{1}{\Delta V_0 A}}, \frac{1}{2\tau_2 \ln \left( \frac{V_{OH} - V_{OL}}{\Delta V_0} \right)} \right\} \]

(3.12)

\[ \Delta V_0 > V_{os, latch} \]

3.3 Operational amplifier

An operational amplifier, which is often called an op-amp, is a DC-coupled high-gain electronic voltage amplifier with differential inputs and single or differential outputs. It is widely used as key block in SHA, MDAC and bandgap voltage reference of pipeline ADC. The notation of ideal op-amp is figure 3.14.
V+ is non-inverting input, V- is inverting input, Vout is op-amp output, \( V_{S+} \) is positive power supply and \( V_{S-} \) is negative power supply. For open loop op-amp \( V_{\text{out}} = (V_+ - V_-)A_{\text{open-loop}} \). \( A_{\text{open-loop}} \) is the open-loop voltage gain of op-amp. The magnitude of the open-loop gain is typically very large, so open-loop operation results in op-amp saturation unless the differential input voltage is extremely small. Op-amps are usually applied to negative-feedback configurations system.

Since practical boundary conditions limit the op-amp performance, such as process specification (Vth, Cox, mobility, etc.), supply voltage and range, operating temperature and range and so on. So op-amp specifications are crucial constraints for the system. These features should be paid attention to:

a. Open loop gain

b. 3-dB bandwidth (dominant pole, Gain bandwidth product)

c. Settling time

d. Offset
e. Slew rate
f. Input common mode range (ICMR)
g. Output-voltage swing
h. Common-mode rejection ratio (CMRR)
i. Power supply rejection ratio (PSRR)
j. Output impedance
k. Noise (Dynamic range=Input swing/Noise)
l. Power consumption

The following section will discussed these configurations with different op-amp architectures.

**3.3.1 Differential Op-amp architecture**

This section only discusses four most popular and fundamental op-amp architectures.

a. Two-stage Op-amp
b. Telescopic Op-amp
c. Folded-cascode Op-amp
d. Gain-boosting Op-amp

**3.3.2 Two-stage Op-amp**
This is the simplest op-amp and it only consists of 8 transistors. The gain of the first stage is often larger than the gain of second stage. Because the higher gain the first stage has, the lower input referred noise and offset the op-amp is. And in order to achieve high output swing, the second stage gain should be smaller. All the specifications of two-stage op-amp are in table 3.3.

\[
G = g_m (r_{o1} \parallel r_{o4}) \cdot g_m (r_{o6} \parallel r_{o7})
\]

<table>
<thead>
<tr>
<th>DC Gain</th>
<th>( G = g_m (r_{o1} \parallel r_{o4}) \cdot g_m (r_{o6} \parallel r_{o7}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dominant pole</td>
<td></td>
</tr>
<tr>
<td>Unity gain bandwidth</td>
<td></td>
</tr>
<tr>
<td>Non-dominant poles</td>
<td></td>
</tr>
</tbody>
</table>
Slew rate \[ SR = \frac{IM_5}{CL} \]

ICMR \[ V_{DS5(sat)} + V_{GS1} \leq V_{IC} \leq V_{DD} - V_{SG3} + V_{th} \]

Output swing \[ V_{out,pp} = V_{DD} - |V_{OD6}| - V_{OD7} \]

Power dissipation \[ P_{diss} = V_{DD} (I_{D5} + I_{D7}) = \frac{1}{2} V_{DD} (g_{m5} V_{OD5} + g_{m7} V_{OD7}) \]

Input referred noise \[ \overline{V}_n \approx 4kT \left( 2 \frac{2}{3g_{m1,2}} + 2 \frac{2g_{m3,4}}{3g_{gm1,2}} \right) + 2 \frac{K_N}{(WL)_{1,2} C_{ox} f} \]
\[ + 2 \frac{K_P}{(WL)_{3,4} C_{ox} f} \frac{g_{m3,4}}{g_{gm1,2}} \]

* \( V_{OD} \) is transistor overdrive voltage. \( V_{OD} = V_{GS} - V_{th} \)

Table 3.3 The Specifications of Two-stage Op-amp.

### 3.3.3 Telescopic Op-amp

![Telescopic Op-amp Diagram](image-url)

Figure 3.16 Telescopic Op-amp.
Telescopic op-amp only has one stage compared with two-stage op-amp. And its gain is as large as two-stage one. However, its output swing is $2^*V_{OD}$ less than that of two-stage one. All the specifications of two-stage op-amp are in table 3.4.

<table>
<thead>
<tr>
<th>DC Gain</th>
<th>$G = g_{m1}[(g_{m3}r_{o3}r_{o1}) \parallel (g_{m5}r_{o5}r_{o7})]$</th>
</tr>
</thead>
</table>

| Dominant pole | $\omega_{-3dB} = p_1 = \frac{-1}{C_L[(g_{m3}r_{o3}r_{o1}) \parallel (g_{m5}r_{o5}r_{o7})]}$ |

| Unity gain bandwidth | $GB = G \cdot \omega_{-3dB} = \frac{g_{m1}}{C_L}$ |

| Non-dominant poles | $p_2 = \frac{-g_{m3}}{C_{gs} + C_{gd1} + C_b}, p_3 = \frac{-g_{m5}}{C_{gs5} + C_{gd7} + C_b5}$ |

| Slew rate | $SR = \frac{I_{ss}}{C_L}$ |

| ICMR | $V_{ISS} + V_{GS1} \leq V_{IC} \leq V_{b1} - V_{GS3} + V_{th1}$ |

| Output swing | $V_{out,pp} = 2[V_{DD} - (V_{OD1} + V_{OD3} + |V_{OD5}| + |V_{OD7}| + V_{ISS})]$ |

| Power dissipation | $P_{diss} = V_{DD}I_{ISS} = \frac{1}{2}V_{DD}(g_{m1}V_{OD1} + g_{m2}V_{OD2})$ |

| Input referred noise | $\overline{V_n}^2 = 4kT \left( \frac{2}{3g_{m1,2}} + 2 \frac{2g_{m7,8}}{3g_{gm1,2}^2} \right) + 2 \frac{K_N}{(WL)_{1,2}C_{ox}f}$ $+ 2 \frac{K_P}{(WL)_{1,2}C_{ox}f} \frac{g_{m7,8}^2}{g_{gm1,2}^2}$ |

* VOD is transistor overdrive voltage. $V_{OD} = V_{GS} - V_{th}$

Table 3.4 The Specifications of Telescopic Op-amp.
### 3.3.4 Folded-cascode Op-amps

Its gain is all most the same as the telescopic one, and the output swing is slightly larger than that of telescope one, $2*V_{OD}$. However the power consumption is twice larger of previous one, because two branches currents run through $V_{DD}$ to ground. One of great features in fold-cascode op-amp is its input signal is larger than other type op-amp. Take figure 3.17 as an example, the lowest input signal could be even less than ground potential. Table 3.5 shows all its specifications.

![Figure 3.17 Folded-cascode Op-amps.](image)

Its gain is all most the same as the telescopic one, and the output swing is slightly larger than that of telescope one, $2*V_{OD}$. However the power consumption is twice larger of previous one, because two branches currents run through $V_{DD}$ to ground. One of great features in fold-cascode op-amp is its input signal is larger than other type op-amp. Take
Figure 3.17 as an example, the lowest input signal could be even less than ground potential. Table 3.5 shows all its specifications.

<table>
<thead>
<tr>
<th>DC Gain</th>
<th>[ G = g_{m1} \left[ \frac{g_{m3} \cdot r_{o3} (r_{09} | r_{01})}{(g_{m5} r_{o5} r_{o7})} \right] ]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dominant pole</td>
<td>[ \omega_{-3dB} = p_1 = \frac{-1}{C_L (g_{m3} r_{o3} (r_{09} | r_{01})) (g_{m5} r_{o5} r_{o7})} ]</td>
</tr>
<tr>
<td>Unity gain bandwidth</td>
<td>[ GB = G \cdot \omega_{-3dB} = \frac{g_{m1}}{C_L} ]</td>
</tr>
<tr>
<td>Non-dominant poles</td>
<td>[ p_2 = \frac{-g_{m3}}{C_{gs1} + C_{gs9} + C_{gs3} + C_{b3}}, p_3 = \frac{-g_{m5}}{C_{gr5} + C_{pr7} + C_{b5}} ]</td>
</tr>
<tr>
<td>Slew rate</td>
<td>[ SR = \frac{I_{ss}}{C_L} ]</td>
</tr>
<tr>
<td>ICMR</td>
<td>[ V_{OD9} \leq</td>
</tr>
<tr>
<td>Output swing</td>
<td>[ V_{out,pp} = 2[V_{DD} - (V_{OD3} + V_{OD9} + V_{OD5}) +</td>
</tr>
<tr>
<td>Power dissipation</td>
<td>[ P_{diss} = V_{DD} (I_{ISS} + I_{M7} + I_{M8}) = V_{DD} (I_{M9} + I_{M10}) = V_{DD} (g_{m1} V_{OD1} + g_{m7} V_{OD7}) ]</td>
</tr>
<tr>
<td>Input referred noise</td>
<td>[ \frac{V_n^2}{2} = 4kT \left( \frac{2}{3g_{m1,2}} + \frac{2g_{m7,8}}{3g_{m1,2}} + \frac{2g_{m9,10}}{3g_{m1,2}} \right) + \frac{K_P}{(WL)<em>{1,2} C</em>{os} f} + \frac{K_P}{(WL)<em>{1,2} C</em>{os} f} \frac{g_{m7,8}^2}{g_{m1,2}^2} + \frac{K_N}{(WL)<em>{1,2} C</em>{os} f} \frac{g_{m9,10}^2}{g_{m1,2}^2} ]</td>
</tr>
</tbody>
</table>

* VOD is transistor overdrive voltage. \[ V_{OD} = V_{GS} - V_{th} \]

Table 3.5 The Specifications of Folded-cascode Op-amp.

52
3.3.5 Gain-boosting Op-amp

The output resistance of figure 3.18(a) is $R_{out} = g_{m2}r_{o2}r_{o1}$, when output of an amplifier is connected to the gate of $M_2$; non-inverting input of the amplifier is set to a fixed voltage $V_b$, and inverting input is connected to the drain of $M_1$. The amplifier boosts the output resistance $R_{out} = A_1g_{m2}r_{o2}r_{o1}$.

![Figure 3.18(a, b) Gain Boosting Technology, (c) Folded-cascode Op-amp with Gain Boosting.](image)

Figure 3.18(a, b) Gain Boosting Technology, (c) Folded-cascode Op-amp with Gain Boosting.
With this boosting technology, the gain of folded-cascode op-amp in Figure 3.18(c) significantly increases [19]. Table 3.6 shows all its specifications. Dominant pole, non-dominant poles, unity gain bandwidth, slew rate, ICMR, output swing is almost the same as folded-cascode one, only introducing more parasitic capacitors.

\[
G = g_m \left[ (A_1 g_{m3} r_{o3} (r_{o9} || r_{o1})) || (A_2 g_{m5} r_{o5} r_{o7}) \right]
\]

| DC Gain | \( G = g_m \left[ (A_1 g_{m3} r_{o3} (r_{o9} || r_{o1})) || (A_2 g_{m5} r_{o5} r_{o7}) \right] \) |
|---------|----------------------------------------------------------------------------------|
| Dominant pole | \( \omega_{3dB} = p_1 = \frac{-1}{C_L \left[ (A_1 g_{m3} r_{o3} (r_{o9} || r_{o1})) || (A_2 g_{m5} r_{o5} r_{o7}) \right] } \) |
| Power dissipation | \( P_{diss} = V_{DD} (I_{ISS} + I_{M7} + I_{M8}) + P_{A1} + P_{A2} = V_{DD} (g_{m1} V_{OD1} + g_{m7} V_{OD7}) + P_{A1} + P_{A2} \) |
| Input referred noise | \( \overline{V_n}^2 = 4kT \left( 2 \frac{2}{3g_{m1,2}} + 2 \frac{2g_{m7,8}}{3g^2_{m1,2}} + 2 \frac{2g_{m9,10}}{3g^2_{m1,2}} \right) \) \\
| | \( + 2 \frac{K_p}{(WL)_{1,2} C_{o9} f} + 2 \frac{K_p}{(WL)_{1,2} C_{o9} f} g^2_{m7,8} + 2 \frac{K_N}{(WL)_{1,2} C_{o9} f} g^2_{m9,10} \) \\
| | \( + \frac{2}{g^2_{m1,2}} + \frac{2}{g^2_{m1,2}} + \frac{2}{g^2_{m1,2} r_{o9}} + \frac{2}{g^2_{m1,2} r_{o9}} \) |

* \( V_{OD} \) is transistor overdrive voltage. \( V_{OD} = V_{GS} - V_{th} \) \( P_{A1} \) and \( P_{A2} \) are power dissipation of additional amplifier \( A_1 \) and \( A_2 \).

Table 3.6 The Specifications of Gain-boosting Op-amp.

Table 3.7 shows the performance comparisons of these four types of op-amps. Two-stage has highest output swing, telescopic one has the fastest speed because of lowest parasitic capacitor; folded-cascode one has the largest ICMR. Gain-boosting op-amp gets the highest DC gain while the circuitry is the most complicated.
### Table 3.7 The Performance Comparisons of Op-amps.

<table>
<thead>
<tr>
<th></th>
<th>DC Gain</th>
<th>Output swing</th>
<th>Speed</th>
<th>Power</th>
<th>Noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-stage</td>
<td>Lowest</td>
<td>Highest</td>
<td>Slowest</td>
<td>Middle</td>
<td>Smallest</td>
</tr>
<tr>
<td>Telescopic</td>
<td>Middle</td>
<td>Smallest</td>
<td>fastest</td>
<td>least</td>
<td>Smaller</td>
</tr>
<tr>
<td>Folded-cascode</td>
<td>Middle</td>
<td>Middle</td>
<td>Middle</td>
<td>Middle</td>
<td>Bigger</td>
</tr>
<tr>
<td>Gain-boosting</td>
<td>Biggest</td>
<td>Middle</td>
<td>Middle</td>
<td>Most</td>
<td>Biggest</td>
</tr>
</tbody>
</table>

### 3.3 SHA AND MDAC

SHA is the most crucial building block in pipeline ADCs. It samples analog signal at one period, and holds the value at the other period. The purpose of this circuit is to hold the analogue value steady for a short time while the converter or other following system performs some operation that takes a little time. MDAC contains a DAC, a subtractor and an S/H amplifier. SHA and MDAC are discussed together, since the structures are the same. The key sub-block of SHA and MDAC is op-amp. Op-amps’ speed, open loop gain, offset will be the limitation of the pipeline ADCs performance.
3.4.1 SHA

Figure 3.19 is the SHA [1, 12]. It works in two non-overlapped phases. The sampling capacitor is C and the parasitic capacitance at the input is pC. The capacitance bC and cC represent the load capacitance switched to the op-amp output on clock phase 2 and 1 respectively. The capacitor aC is the feedback capacitor to make an accurate SHA gain.

![SHA Diagram](image)

Figure 3.19 (a) SHA (b) Phase 1 (c) Phase 2.
Op-amp has input offset \(V_{OS}\) and the finite open loop gain \(A\) which will impact the overall performance of SHA. At phase 1, the analog signal \(V_{in}\) samples on the capacitor \(C\), the time constant is: \(\tau_{\text{sampling}} = (R_{ON} + R_s)C\), and voltage on capacitor \(C\) is

\[
\tau_{\text{sampling}} = (R_{ON} + R_s)CV_{cap} = V_{in}(1 - e^{-t/\tau_{\text{sampling}}}) \tag{3.13}
\]

The relative settling error should be:

\[
\epsilon_{rP1} = \frac{V_{in} - V_{cap}}{V_{in}} = e^{-t/\tau_{\text{sampling}}} \tag{3.14}
\]

The charge in \(C\) and \(pC\) are:

\[
Q_{p1} = -CV_{in} + \frac{AV_{OS}}{A+1}pC \tag{3.15}
\]

At phase 2, using small equivalent signal model, the transform function of SHA is:

\[
\frac{V_{OUT}}{V_{IN}} = \frac{-C(g_m - aCs)R_O}{aCR_Og_m + C + pC + aC + R_O[bC(C + aC + pC) + (C + pC)a]s} \approx \frac{-(g_m - aCs)R_O}{aR_Og_m + R_O[b(C + aC + pC) + (C + pC)a]s} \tag{3.16}
\]

Time constant is \(\tau_{\text{amp}} = \frac{C(a+b)(1+p)+ab}{g_m}a\), and the dynamic relative settling error:

\[
\epsilon_{rP2} = \frac{V_{in} - V_{cap}}{V_{in}} = e^{-t/\tau_{\text{amp}}} \tag{3.17}
\]
The charge at negative input terminal:

\[ Q_{p2} = (C + pC)V_{{n}} + aC(V_{{n}} - V_{{OUT}}) = (1 + p + a)CV_{{n}} - aCV_{{OUT}} = (1 + p + a)C\left( V_{{OS}} - \frac{V_{{OUT}}}{A} \right) - aCV_{{OUT}} \]  

\[ (3.18) \]

The charges at that node during phase1 and phase2 should be the same:

\[ Q_{p2} = Q_{p1} \implies \]

\[ (1 + p + a)C\left( V_{{OS}} - \frac{V_{{OUT}}}{A} \right) - aCV_{{OUT}} = -CV_{{in}} + \frac{AV_{{OS}}}{A+1}pC \]  

\[ (3.19) \]

\[ \implies V_{{OUT}} = \frac{V_{{in}} + (1 + a + p)V_{{OS}} - \frac{AV_{{OS}}}{A+1}p}{a + \frac{1 + a + p}{A}} \]

In order to achieve an accurate value of SHA, resolution and sampling rate requirements should be trade-off: the total errors consist of static error and dynamic error. The finite gain and input offset of op-amp and capacitance mismatch result in static error; for dynamic one, op-amp takes a certain time to reach the final value.

The worst error case: \( V_{{in}} = V_{{ref}} \), total error should be less than LSB.

- At phase1, both dynamic error and input referred noise will limit the performance

Dynamic error:

\[ \varepsilon_{{total}} \approx \varepsilon_{{dynamic}} = \varepsilon_{{t^1}} = e^{-t/\tau_{{sampling}}} < \frac{1}{2^N} \]

\[ \implies t_1 > \tau_{{sampling}} \ln(2^N) = N\tau_{{sampling}} \ln 2 = N(R_{{ON}} + R_{{s}})C \ln 2 \]  

\[ (3.20) \]
Dynamic range:

- Input referred noise: $V_n^2 = \frac{kT}{C}$ 
  \hspace{1cm} (3.21)

- If a sinusoidal signal is applied:

$$DR = \frac{V_{\text{input}}^2}{V_n^2} = \frac{1}{2} \frac{V_{\text{input,swing}}^2}{kT/C}$$  \hspace{1cm} (3.22)

- ADC should make sure SNR<DR, it means all the valid input signal will be converted by the ADC, then:

$$\frac{1}{2} \left( \frac{V_{\text{input,swing}}}{2} \right)^2 > 1.5 \cdot 2^N \Rightarrow C > \frac{12 \cdot 2^N kT}{V_{\text{input,swing}}^2}$$  \hspace{1cm} (3.23)

- At phase 2, both static and dynamic errors impact the performance. For partition error:

$$\varepsilon_{\text{dynamic}} < \frac{1}{2} \varepsilon_{\text{total}} = \frac{1}{2^{N+1}}, \quad \varepsilon_{\text{static}} < \frac{1}{2} \varepsilon_{\text{total}} = \frac{1}{2^{N+1}}$$  \hspace{1cm} (3.24)

- Dynamic error:
\[ E_{\text{dynamic}} \approx E_{r2} = e^{-t/\tau_{\text{amp}}} < \frac{1}{2^{N+1}} \]

\[ t_2 > \tau_{\text{amp}} \ln(2^{N+1}) = (N + 1)\tau_{\text{amp}} \ln 2 = (N + 1) \frac{C}{g_m} \frac{(a + b)(1 + p) + ab}{a} \ln 2 \]

\[ \Rightarrow t_2 > (N + 1) \frac{C}{g_m} \frac{(a + b) + ab}{a} \ln 2 \]

- Static errors: op-amp offset, finite gain, input parasitic capacitance and feedback capacitance mismatch all impact the static error.
  - For perfect capacitance layout matching, \( a = 1 \pm 0.1\% \);
  - Parasitic cap. \( C_{\text{gs}} = 2/3(WL)C_{\text{ox}} \)
  - The typical VOS without offset cancellation is 5mV~30mV.
  - Only if considering finite gain \( A \), when \( V_{\text{in}} = V_{\text{ref}} \), the worst case:

\[ V_{\text{out}} \approx \frac{V_{\text{in}} + (2 + \Delta a)V_{\text{os}}}{1 + \left( \frac{\Delta a + 2 + \Delta a + p}{A} \right)} \approx V_{\text{in}} \left( 1 - \Delta a - \frac{2 + \Delta a}{A} \right) + 2V_{\text{os}} \]

\[ \left| \frac{\Delta a + 2 + \Delta a}{A} \right| V_{\text{ref}} < \text{LSB} \Rightarrow \left| \frac{\Delta a + 2 + \Delta a}{A} \right| V_{\text{ref}} < \frac{2V_{\text{ref}}}{2^N} \]

\[ \Rightarrow A > \frac{2 + \Delta a}{1 - \Delta a} \]

\[ \frac{1}{2^{N-1}} > \Delta a > -\frac{2}{A + 1} \Rightarrow \frac{1}{2^{N-1}} - \Delta a > 0 \Rightarrow N < -\frac{\Delta a}{\lg 2} + 1 \]
3.4.2 MDAC

Figure 3.20 is the MDAC [1]. It works in two non-overlapped phases. The sampling capacitor is C and the parasitic capacitance at the input is pC. The capacitance bC and cC represent the load capacitance switched to the op-amp output on clock phase 2 and 1 respectively. The capacitor aC is the feedback capacitor to make an accurate MDAC gain.
With the same method of SHA, at the phase 1, the input signal samples on capacitor C and aC:

\[
\tau_{\text{sampling}} = (R_{ON} + R_S)(C + aC)
\]
\[
\tau_{rP1} = e^{-t_1/\tau_{\text{sampling}}}
\]
\[
Q_{p1} = -(1 + a)CV_{in} + \frac{AV_{OS}}{A+1} pC
\]  (3.27)

At the phase 2, the circuit is exactly the same as SHA in its phase2:

\[
Q_{p2} = (1 + p + a)C(V_{OS} - \frac{V_{OUT}}{A}) - aCV_{OUT}
\]
\[
\tau_{\text{amp}} = \frac{C}{g_m} \frac{(a+b)(1+p)+ab}{a}, \tau_{rP2} = e^{-t_2/\tau_{\text{amp}}}
\]  (3.28)

The charges during phase 1 and phase 2 should be the same:
\[ Q_{f_2} = Q_{f_1} \quad \Rightarrow (1 + p + a)C(V_{os} - \frac{V_{OUT}}{A}) - aCV_{OUT} = -(1 + a)CV_{in} + \frac{AV_{os}}{A+1}pC \]

\[ \Rightarrow V_{OUT} = \frac{(1 + a)V_{in} + (1 + a + p)V_{os} - \frac{AV_{os}}{A+1}p}{a + \frac{1 + a + p}{A}} \]

\[ \approx \frac{(1 + a)V_{in} + (1 + a)V_{os}}{a + \frac{1 + a}{A}} \approx \frac{(1 + \frac{1}{a})V_{in} + (1 + \frac{1}{a})V_{os}}{1 + \frac{1}{a}} \approx (1 + \frac{1}{a})(V_{in} + V_{os}) \left( \frac{1 + \frac{1}{a}}{A} \right) \]

\[ \approx \left( \frac{1}{a} + 1 \right)V_{in} \left( 1 - \frac{1}{a} \right) + (1 + \frac{1}{a})V_{os} \]

\[ \left( \frac{1}{a} + 1 \right) \frac{1 + \frac{1}{a}}{A} V_{ref} < \text{LSB}_{left} = \frac{2V_{ref}}{2^{N_{ref}}} \Rightarrow A > 2^{N_{ref}-1} \left( 1 + \frac{1}{a} \right)^2 \]

(3.29)

Reconsideration performance limitations in MDAC again:

\[ \frac{1}{2f_S} > t_1 > \tau_{\text{sampling}} \quad \ln(2^{N_{ref}}) = N_{left} \tau_{\text{sampling}} \ln 2 = N_{left} (R_{ON} + R_s) C \ln 2 \]

\[ \varepsilon_{\text{dynamic}} < \frac{1}{2} \varepsilon_{\text{total}} = \frac{1}{2^{N_{ref}+1}}, \quad \varepsilon_{\text{static}} < \frac{1}{2} \varepsilon_{\text{total}} = \frac{1}{2^{N_{ref}+1}} \]

\[ \frac{1}{2f_S} > t_2 > \tau_{\text{amp}} \ln(2^{N_{ref}+1}) = (N_{left} + 1) \tau_{\text{amp}} \ln 2 = (N_{left} + 1) \frac{C}{g_m} (a + b)(1 + p) + ab \ln 2 \]

\[ A > 2^{N_{ref}-1} \left( 1 + \frac{1}{a} \right)^2 \]

\[ C > \frac{3 \cdot kT}{V_{\text{input swing}}^2} 2^{2N-2} = \frac{1}{4} C_{\text{previous}} \]

(3.30)
3.5 Voltage reference

Voltage Reference is a circuit used to generate a fixed voltage, Vref, which is independent of the power supply voltage VDD, temperature, and process variations. There are three types of voltage references in figure 3.21:

a. Proportional to absolute temperature (PTAT)
b. Complementary to absolute temperature (CTAT)
c. Very little changes with temperature (BGR)

![Figure 3.21 Voltage Reference.](attachment://figure_3_21.png)

Two of most important voltage Reference Specifications are temperature coefficient and PSRR.

a. Temperature coefficient The unit is ppm/°C = 10-6/°C (parts per million per degree C)
b. Power supply rejection ratio (PSRR): describe the amount of noise from a power supply that the voltage reference can reject.
The smaller the temperature coefficient voltage reference is, the stable and accurate the system it is. And PSRR should be as large as possible, to minimize the voltage supply interference, especially in the mixed-signal systems, the analog voltage supply is noisy.

### 3.5.1 Bandgap reference

Bandgap reference circuits are necessarily and widely used in data-conversion systems, voltage regulators and memories. They can provide very stable references hardly dependent on temperature and external power supply.

![Bandgap Voltage Reference](image)

Figure 3.22 Bandgap Voltage Reference.

The basic idea of BGR is to add a proportional to absolute temperature (PTAT) voltage to the emitter-base voltage (VBE), so that the first-order temperature dependency
of the pn junction is compensated by the PTAT voltage, and a nearly temperature
independent output is generated. The PTAT voltage is actually the thermal voltage (\(V_t\))
of the pn junction [28, 29, 30]. Thus the reference voltage is expressed as:

\[
V_{R1} = V_{BE1} - V_{BE2} = \Delta V_{BE} \implies I_{R1} = \frac{\Delta V_{BE}}{R_1} = \frac{V_T \ln n}{R_1}
\]

\[
V_{ref} = V_{BE2} + I_{R1}(R_1 + R_2) = V_{BE3} + \frac{V_T \ln n}{R_1}(R_1 + R_2)
\]

\[
= V_{BE2} + V_T (1 + L) \ln n
\]

\[
\frac{\partial V_{ref}}{\partial T} = \frac{\partial V_{BE2}}{\partial T} + (1 + L) \cdot \frac{k}{q} \ln n = 0
\]

\[
\frac{\partial V_{BE3}}{\partial T} \approx -1.6mV/C
\]

### 3.5.2 Curvature-compensated BGR

The more accurate emitter-base voltage (VBE) of BJT is:

\[
V_{BE}(T) = V_{BG} - (V_{BG} - V_{BE0}) \frac{T}{T_0} - (\eta - \alpha) V_T \ln \frac{T}{T_0}, \quad \eta \approx 4
\]

- \(\alpha = 0\), when the current in BJT is PTAT.
- \(\alpha = 1\), when the current in BJT is temperature independent.

Because of logarithm function in that equation, higher order temperature coefficients
exist. The previous BGR method only cancels the first order temperature coefficient.
Curvature-compensated BGR in figure 3.5.3 cancels higher order temperature coefficient
[33, 34, 35].
\[ V_{BEQ3}(T) = V_{BG} - (V_{BG} - V_{BE0}) \frac{T}{T_0} - \eta V_T \ln \frac{T}{T_0} \]

\[ V_{BEQ1,2}(T) = V_{BG} - (V_{BG} - V_{BE0}) \frac{T}{T_0} - (\eta - 1) V_T \ln \frac{T}{T_0} \]

\[ V_{NL} \approx V_{BEQ3}(T) - V_{BEQ1,2}(T) = V_T \ln \frac{T}{T_0} \] (3.34)

Figure 3.23 Curvature-compensated BGR.

Curvature compensation can be achieved by subtracting from both $I_1$ and $I_2$ a current proportional to $V_{NL}$, where $V_T = kT/q$ and $N$ is emitter area ratio of transistor $Q_1$ and $Q_2$. In the equation above, $VTLnN/R_1$ is PTAT current compensating the first-order temperature dependency of $VEBQ2/R_2$. Therefore, sum of $I_1$ and $I_2$ is nearly temperature independent. Besides, since emitter current of $Q_1$ and $Q_3$ are PTAT and temperature-independent respectively, $(VEBQ2-VEBQ3)/R_5$ provides current to compensate higher-order temperature dependency in $VEBQ2/R_2$. 
\[ V_{out} = R_3 \left( \frac{1}{R_1} V_{BE1} + \frac{V_T \ln n}{R_0} + V_{NL} \left( \frac{1}{R_{4,5}} \right) \right) \]

\[ = \frac{R_3}{R_1} \left( V_{BE1} + \frac{R_1 \ln n}{R_0} V_T + V_{NL} \left( \frac{R_1}{R_{4,5}} \right) \right), R_{4,5} = \frac{R_1}{\eta - 1} \quad (3.35) \]

### 3.5.3 Bandgap voltage reference in pipeline ADC

BGR does not only supply accurate voltage reference in ADC, but also generates biasing circuits. Since voltage reference is more accurate and crucial, biasing circuits in pipeline ADC associated with BGR are omitted. The higher the resolution of ADC is, the more accurate BGR is required. Because more decision levels are needed, so the voltage reference should be more detailed. Let’s consider PSRR and Temperature coefficient impaction on nonlinearity and accuracy of ADC.

#### 3.5.3.1 Nonlinearity

The reference generator (Figure 3.24a) in ADC is implemented by BGR. It compares the input signal with the reference by comparators. The simplest reference generator is figure 3.24(b). Supposing all the resistors are ideal and don’t have mismatching or variations.
Figure 3.24 Reference Generator.

$V_H$ is the highest possible input signal value and $V_L$ is the lowest possible input signal value. And BGR generate a voltage reference $V_{\text{ref}}$.

\[
\begin{align*}
\alpha_r &= \frac{R_r}{R_{\text{tot}}} , \alpha_A = \frac{R_A}{R_{\text{tot}}} \\
V_H &= \frac{R_r + R_A}{R_{\text{tot}}} V_{\text{ref}} = (\alpha_r + \alpha_A) \times V_{\text{ref}} \\
V_L &= \frac{R_A}{R_{\text{tot}}} V_{\text{ref}} = \alpha_A \times V_{\text{ref}} \\
V_r &= V_H - V_L = \alpha_r \times V_{\text{ref}} \\
\Delta &= \frac{V_r}{2^n} = \frac{\alpha_r}{2^n} V_{\text{ref}}
\end{align*}
\]

(3.36)

Then all $(2n-1)$’s the decision levels are:
For $1 \leq i \leq 2^n - 1$

$$V_i = V_0 + i \Delta = \alpha_A \cdot V_{ref} + i \frac{\alpha_r}{2^n} V_{ref} = \left(\alpha_A + \frac{\alpha_r \times i}{2^n}\right) \times V_{ref}$$  \hspace{1cm} (3.37)$$

$$V_{i+1} = V_0 + (i + 1) \Delta = \left(\alpha_A + \frac{\alpha_r \times (i + 1)}{2^n}\right) \times V_{ref}$$

- **DNL**

$$DNL_i = \frac{V_{i+1} - V_i - \Delta}{\Delta}$$

  - The ideal ADC $DNL_i = 0$ $(0 < i < 2n)$
  - Practical ADC: voltage reference varies. $\Delta V_{ref}$ is the variations of BGR.

$V_{ref}$ is the ideal value.

$$V_{ref}' = V_{ref} + \Delta V_{ref}$$

$$DNL_i = \frac{V_{i+1}' - V_i' - \Delta}{\Delta} = \frac{\left(\alpha_A + \frac{\alpha_r \times (i + 1)}{2^n}\right) \times V_{ref}' - \left(\alpha_A + \frac{\alpha_r \times i}{2^n}\right) \times V_{ref}' - \Delta}{\Delta}$$

$$DNL_i = \frac{\alpha_r \times V_{ref}' - \alpha_r \times V_{ref}}{\Delta} = \frac{\Delta V_{ref}}{V_{ref}}$$

$$= \frac{\alpha_r \times V_{ref}}{2^n} \frac{\Delta V_{ref}}{V_{ref}} = \frac{\Delta V_{ref}}{V_{ref}}$$  \hspace{1cm} (3.38)$$

- **INL**

$$INL_i = \frac{V_i' - V_i}{\Delta} = \frac{\left(\alpha_A + \frac{i}{2^n}\right) \times V_{ref}' - \left(\alpha_A + \frac{i}{2^n}\right) \times V_{ref}}{\Delta}$$

$$= \frac{\left(\alpha_A + \frac{i}{2^n}\right) \Delta V_{ref}}{\Delta} = \frac{\alpha_A}{\alpha_r} \frac{2^n + i}{V_{ref}} \Delta V_{ref}$$  \hspace{1cm} (3.39)$$

  - When $i=2n-1$, $INL_i$ is maximum.
\[ \text{INL}_{\text{max}} = \left( \frac{\alpha_A}{\alpha_r} \right) \frac{2^n + 2^n - 1}{V_{\text{ref}}} \Delta V_{\text{ref}} \] (3.40)

In order to get no missing code, the requirement of DNL and INL should be:

\[ |\text{INL}_i| \leq \frac{1}{2}, \text{ for all } i; \quad |\text{DNL}_i| \leq 1, \text{ for all } i \] (3.41)

BGR variations requirement should be:

\[
\begin{align*}
\text{INL}_{\text{max}} &= \left( \frac{\alpha_A}{\alpha_r} \right) \frac{2^n + 2^n - 1}{V_{\text{ref}}} \Delta V_{\text{ref}} \leq \frac{1}{2} \\
\text{DNL}_i &= \frac{\Delta V_{\text{ref}}}{V_{\text{ref}}} \leq 1 \\
\Rightarrow \Delta V_{\text{ref}} &\leq \min \left[ V_{\text{ref}}, \frac{V_{\text{ref}}}{2 \left( \frac{\alpha_A}{\alpha_r} \right) 2^n + 2^n - 1} \right]
\end{align*}
\] (3.42)

3.5.3.2 Accuracy

- **PSRR:**

The voltage reference should be varied less than half of 1 LSB; otherwise, wrong voltage comparisons are made. Let’s suppose the voltage supply is not stable, varying by \( \Delta V_{\text{DD}} \), So:

\[
\Rightarrow \text{PSRR} = 20 \cdot \log \left( \frac{\Delta V_{\text{DD}}}{V_{\text{ref}}} \right) \geq 20 \cdot \log \left( \frac{\Delta V_{\text{DD}}}{V_{\text{ref}}} \right) 2 \left( \frac{\alpha_A}{\alpha_r} \right) 2^n + 2^n - 1 \]

\[
> 20 \cdot \log \left( \frac{\Delta V_{\text{DD}}}{V_{\text{ref}}} 2^{n+1} \left( \frac{\alpha_A}{\alpha_r} + 1 \right) \right) = 6.02(n + 1) + 20 \log \left( \frac{\Delta V_{\text{DD}}}{V_{\text{ref}}} \right) + 20 \log \left( \frac{\alpha_A}{\alpha_r} + 1 \right)
\] (3.43)
o For example,

- n=16bit, Vref=2.4V, VDD=3.3V, VL=1.1V, VH=2.225V

\[
\alpha_a = \frac{1.1}{2.4}, \alpha_r = \frac{1.125}{2.4}, \Delta VDD = 0.25V
\]

- PSRR=6.02*17-20+6=88.6dB

- If \( \Delta VDD=10mV \), PSRR=60.67dB

- The voltage temperature coefficient:

The circuit temperature varies during working, which results in the voltage reference variations.

\[
TCV_{ref} = \frac{\Delta V_{ref}}{V_{ref} \Delta T} \leq \frac{1}{2 \left( \frac{\alpha_a}{\alpha_r} \right) \left( 2^n + 2^n - 1 \right) \Delta T} < \frac{1}{2^{n+1} \left( \frac{\alpha_a}{\alpha_r} + 1 \right) \Delta T}
\] (3.44)

- For example,

  o n=16bit, Vref=2.4V, VDD=3.3V, VL=1.1V, VH=2.225V

\[
\alpha_a = \frac{1.1}{2.4}, \alpha_r = \frac{1.125}{2.4}
\]

\[
TCV_{ref} = \frac{1}{2^{n+1} \left( \frac{\alpha_a}{\alpha_r} + 1 \right) \Delta T} \approx \frac{2^{-18}}{\Delta T} \approx 3.814 \times 10^{-6}/\Delta T
\]
CHAPTER 4

PIPELINE ADC POWER OPTIMIZATION

Typical pipeline ADCs categorized by single-bit and multi-bit per stage perform analog-to-digital conversion process in several stages following an SHA. Each stage is built up with one sub-flash ADC and one MDAC consisting of a switch capacitor DAC and a closed-loop residue amplifier. Three design issues – speed, resolution and power, constitute the performance matrix. The speed is often decided by the settling time of the sample and hold circuit and the residue amplifiers. Besides the non-full settling, capacitor mismatch and gain error due to the finite DC gain of the residue amplifier contribute to the conversion accuracy of data converters. Most of the power dissipation comes from the sample and hold circuit and the residue amplifiers. The potential optimization points in traditional pipelined ADCs are the speed limitation generated from the switch capacitor based closed-loop residue amplifiers, the accurate output range requirement on each stage due to the fixed reference ladder and the capacitor mismatch constrained by the process.

In this section, one optimized pipeline ADC method is proposed to release these design limitation by stage partition and op-amp power to comparator power ratio. The feasibility and performance improvement have been analyzed in detail. The major power dissipation components in each stage of pipeline ADC will be sub-ADC and MDAC. Biasing circuits, clock generators and digital circuitry are much smaller. So, it’s
reasonable to neglect these power dissipations. Before structure optimization, some denotation should be clear:

- N as the total number of bits
- B as the number of bits/stage
- n as the number of stages
- F as the normalized total power
- R as ratio between the power consumed by MDAC and that by the comparator

### 4.1 Pipeline with identical stages

For pipeline ADC with identical stages, the same capacitors are used for all MDACs. To maintain the same speed, the same transconductance is required for the amplifier in each MDAC. So, each MDAC consumes the same power [21].

The total power consumption of the ADC normalized to that of the comparator is derived to be:

\[
F = \frac{P_{\text{total}}}{P_{\text{comp}}} = \frac{n \times P_{\text{MDAC}} + n \times (2^B - 2) \times P_{\text{comp}}}{P_{\text{comp}}} = \frac{n \times P_{\text{MDAC}}}{P_{\text{comp}}} + n \times (2^B - 2)
\]

\[
= n \times \left[ R + (2^B - 2) \right] = \frac{N - 1}{B - 1} \left[ R + (2^B - 2) \right]
\]
Figure 4.1 indicates that the total power $F$ increases much faster with the power ratio $R$ for lower $B$ because the number of stages and the number of MDACs required are much larger.

The global minimum value of $F$ can be achieved for $B=2$ and $R=1$, which is not practical.

Differentiating $F$ and solving for $B_{opt}$:

$$F = \frac{N-1}{B-1} \left[ R + \left( 2^B - 2 \right) \right]$$

$$\frac{\partial F}{\partial B} = \frac{N-1}{B-1} 2^B \ln 2 - \frac{N-1}{(B-1)^2} \left[ R + \left( 2^B - 2 \right) \right]$$

$$= \frac{N-1}{(B-1)^2} \left\{ 2^B \left[ (B-1) \ln 2 - 1 \right] - R - 2 \right\} = 0$$

$$\Rightarrow R = 2 - 2^B \left[ 1 - (B-1) \ln 2 \right]$$

(4.2)
We can get the optimized R as figure 4.2. For every given R, we can get the fixed optimized B. Given a power ratio R, $B_{opt}$ is independent of the total number of bits N. $F_{min}$ only depends on the minimum power dissipated by each stage, which is dependent on $B_{opt}$ but independent of N [21].

4.2 Pipeline with capacitor scaling

In pipeline ADC, thermal noise $kT/C$ contribution of a later stage is effectively attenuated by the gain of the previous stages.

$$V_{\text{total}}^2 = V_{n1}^2 + \frac{V_{n2}^2}{A_1^2} + \ldots + \frac{V_{nn}^2}{A_1^2 A_2^2 \ldots A_{n-1}^2}$$

(4.3)

- $\overline{V_{ni}}$ is the total input noise of the i-th stage.
• Ai is the i-th stage residue gain.

So the capacitor of later stages can be scaled down to reduce their power dissipation without increasing the kT/C noise significantly. In the figure 4.3, the capacitors are scaled down by a factor of S/stage, the power of MDAC can be reduced by the same factor without sacrificing the speed.

\[
P_{MDAC} \approx P_{amp} = V_{DD} \times (k \times I_D)
\]

\[
I_D = \frac{1}{2} g_m V_{OD}
\]

\[
g_m = 2\pi \times GWB \times C_L
\]

\[
V_{OD} \approx 5\% \times V_{DD}
\]

\[
\Rightarrow P_{MDAC} \approx V_{DD} \times k \times \frac{1}{2} V_{OD} \times 2\pi \times GWB \times C_L = V_{DD} \times k \times V_{OD} \times \pi \times GWB \times C_L
\]

\[
= 0.05k\pi \times V_{DD}^2 \times GWB \times C_L
\]

As the capacitor scales down by a factor of S, the power consumption is also decrease by the factor of S.
For \( C_{L(i+1)} = \frac{C_L}{S} \Rightarrow P_{MDAC(i+1)} = \frac{P_{MDAC(i)}}{S} \)

\[ P_{total} = P_{MDAC0} + P_{MDAC1} + \ldots + P_{MDAC(n)} + n \times (2^B - 2) \times P_{comp} \]

\[ F = \frac{P_{total}}{P_{comp}} = \frac{\sum_{i=1}^{n} P_{MDAC(i)} + n \times (2^B - 2) \times P_{comp}}{P_{comp}} = \frac{1 - S^{-n}}{1 - S^{-1}} P_{MDAC0} + n \times (2^B - 2) \]

\[ = \frac{1 - S}{1 - S^{-1}} R + \frac{N - 1}{B - 1} \times (2^B - 2) \]  

(4.5)

a. For low power design, S should be larger;

b. For low noise design, S should be smaller.

a) If \( S = G_i = 2^{2(B-1)} \),

\[ \max \left[ \frac{V_{\text{total}}^2}{n_{\text{total}}} \right] = n \times V_{n1}^2, \min[F] = \frac{1 - 2^{-2(N-1)}}{1 - 2^{-2(B-1)}} R + \frac{N - 1}{B - 1} \times (2^B - 2) \]  

(4.6)

b) If \( S = 1 \)

\[ \min \left[ \frac{V_{\text{total}}^2}{n_{\text{total}}} \right] = V_{n1}^2 \times \frac{1 - 2^{-2(N-1)}}{1 - 2^{-2(B-1)}}, \max[F] = \frac{N - 1}{B - 1} R + \frac{N - 1}{B - 1} \times (2^B - 2) \]  

(4.7)

The total power for capacitor scaling is always much smaller compared with identical stage case. For a low power ratio \( R \), the total power \( F \) for the capacitor scaling case increases exponentially. The power contributed by MDACs is negligibly small and that the total power is basically from the sub-ADCs which depend exponentially on \( B \) [21].
Figure 4.4 illuminates that $F$ varies linearly with $R$ for a given $N$ and $B$ for both identical stages and capacitor scaling. $F$ is less sensitive to the total number of bits $N$ for low values of $B$. (With capacitor scaling, the power consumption of the first MDAC is dominant and that of the last few stages can be neglected. Increasing $N$ and the number of stages $n$ only increase the total power slightly).

$$F = \frac{1 - 2^{\frac{N-1}{B-1}}}{1 - S^{-1}} R + \frac{N-1}{B-1} \times (2^B - 2)$$

If $S = 2^{B-1}$, $F = \frac{1 - 2^{-(N-1)}}{1 - 2^{-(B-1)}} R + \frac{N-1}{B-1} \times (2^B - 2)$

$$\frac{\partial F}{\partial B} = -\frac{1 - 2^{-(N-1)}}{\left[1 - 2^{-(B-1)}\right]^2} R \times 2^{-(B-1)} \ln 2 - \frac{N-1}{(B-1)^2} \times (2^B - 2) + \frac{N-1}{B-1} \times 2^B \ln 2 = 0$$

(4.8)

$$\Rightarrow R = \frac{N-1}{B-1} \left[1 - 2^{-(B-1)}\right]^2 \times \left[2^B \ln 2 \times \frac{(2^B - 2)}{B-1}\right]$$

Figure 4.5 tells that $B_{opt}$ for capacitor scaling is dependent on the total number of bit $N$. the flat regions extend over wider ranges of the $R$ because the power consumed by
MDACs are much smaller for later stages. As a result, it would require a much larger increase in $R$ to compensate for an increase in $B_{opt}$.

![Figure 4.5 Optimized B vs. R.](image)

### 4.3 Pipeline with resolution scaling

Capacitor scaling only minimizes the power consumption of MDACs and does not affect the power dissipation of sub-ADC. A scheme with no uniform bits/stage is considered to reduce the power contribution of sub-ADC to the whole system.
\[ F = \frac{P_{\text{total}}}{P_{\text{comp}}} = nR + \sum_{i=1}^{n} (2^{B_i} - 2) = nR + \sum_{i=1}^{n} 2^{B_i} - \sum_{i=1}^{n} 2 \]

\[ = nR + \sum_{i=1}^{n} 2^{B_i} - 2n = n(R - 2) + \sum_{i=1}^{n} 2^{B_i} \]

\[ n = (B_1 + B_2 + \ldots + B_n + 1) - N \tag{4.9} \]

For N=16 bits pipeline ADC, the possible stage partition as follows:

1. 5-5-5-4
2. 4-4-4-4-4
3. 5-5-4-3-3
4. 5-4-4-4-3
5. 4-4-4-3-3-3
6. 4-3-3-3-3-3-3
7. 4-4-3-3-3-3-2
8. 4-4-4-3-3-2-2
9. 3-3-3-3-3-3-3-2
10. 4-3-2-2-2-2-2-2-2-2-2-2-2-2
All the possible power consumption curves are drawn in figure 4.6. None of the bit patterns can yield a global minimum power for all power ratio R. From a low value of R, one pattern may achieve the local minimum power. As R becomes large enough, another pattern will take over. Both resolution scaling and capacitor scaling should be considered. $S = 2^{Bi-1}$
\[
P_{\text{total}} = \sum_{i=1}^{n} P_{MDAC(i)} + \sum_{i=1}^{n} P_{\text{Sub-ADC}(i)}
\]
\[
= \left( P_{MDAC1} + P_{MDAC2} + \ldots + P_{MDACn} \right) + \left( 2^{B_1} - 2 \right) + \left( 2^{B_2} - 2 \right) + \ldots + \left( 2^{B_n} - 2 \right) \times P_{\text{comp}}
\]
\[
= \left( P_{MDAC} + \frac{P_{MDAC}}{2^{B_1 - 1}} + \frac{P_{MDAC}}{2^{B_2 - 1}} \times 2^{B_1 - 1} + \ldots + \frac{P_{MDAC}}{\prod_{i=2}^{n} 2^{B_{i-1}}} \right) + \left[ \sum_{i=1}^{n} 2^{B_i} - 2n \right] \times P_{\text{comp}} \quad (4.10)
\]
\[
F = \frac{P_{\text{total}}}{P_{\text{comp}}} = R \left\{ 1 + \frac{2}{2^{B_1}} + \frac{2^2}{2^{B_2} + 2} + \ldots + \frac{2^{n-1}}{\prod_{i=2}^{n} 2^{B_{i-1}}} \right\} + \left( \sum_{i=1}^{n} 2^{B_i} - 2n \right)
\]
\[
n = (B_1 + B_2 + \ldots + B_n + 1) - N
\]

Figure 4.7 Optimized Power with Both Capacitor and Resolution Scaling vs. R.

Using different scaling factor S, the global minimum power dissipation would be quite different. In practice, due to technology and layout constraints, a minimum value of capacitor that can be used may exist [21]. Minimizing the total chip area may put an upper limit on the maximum capacitor value that can be designed. So, the capacitors for
the first few and last few stages may need to remain un-scaled, and the scaling scheme would not be as effective and would need to be modified.

4.4 Pipeline ADC power optimization with thermal noise

This section shows how to find optimized scaling factor and number of stages with power and thermal noise limitation with ADC specification. Figure 4.8 shows the pipeline ADC structures. Suppose the stages are identical expect the capacitance scaling. This means each stage has the same digital output, has the same op-amp structure but the power consumption might not be the same because of the capacitance scaling. On one hand, if N-bits pipeline ADC has only one stage, the parasitic capacitance of comparators is huge because this stage has $2^N-1$ comparators.

![Diagram of Pipeline ADC Structures](image)

Figure 4.8 Pipeline ADC Structures.

As the N is 8~16 bits, the number of comparator is dominant. However, on the other hand, if each stage only has one bit digital output, N bit pipeline ADC need N
stages, every stage will have one op-amp, which consume lots of power. So, there must be same optimized partition for pipeline ADC, the number of stage must sit between 1 and N.

In pipeline ADC, thermal noise $kT/C$ contribution of a later stage is effectively attenuated by the gain of the previous stages.

$$V_{\text{in}_{\text{total}}}^2 = V_{n1}^2 + \frac{V_{n2}^2}{A_1^2} + \ldots + \frac{V_{nn}^2}{A_1^2 A_2^2 \ldots A_{n-1}^2}$$

(4.11)

$\overline{V}_i$ is the total input noise of the i-th stage, $A_i$ is the i-th stage residue gain. The capacitor of later stages can be scaled down to reduce their power dissipation without increasing the $kT/C$ noise significantly. In the left figure, the capacitors are scaled down by a factor of $S$/stage; the power of MDAC can be reduced by the same factor without sacrificing the speed. The follow is some definition of ADC. $A_C$ is the close-loop gain; $f$ is the feedback factor of op-amps. $M$ is the number of stage. $C_{s,k}$ is the sampling capacitor of $k^{\text{th}}$ stage; $C_{L,k}$ is the total load capacitor of $k^{\text{th}}$ stage; $C_{\text{flash}}$ is the total comparators capacitor of each stage. $s$ is the scaling factor of pipeline ADC.[3]

$$A_C = 2^{n_s - 1} \quad f = \frac{1}{A_c} \quad M = \left\lfloor \frac{N-1}{n_s-1} \right\rfloor$$

$$C_{s,k} = (A_C - 1)C_{F,k} \quad C_{L,k} = A_c C_{F,k+1} + C_{\text{flash}} \quad C_{\text{flash}} = 4A_c (A_C - 1)C_{\text{unit}} \quad a = \frac{C_{\text{unit}}}{C_{s0}}$$

(4.12)
In MDAC, the most important noise sources are sampling switches and thermal noise in the op-amp. In the analysis we ignore flicker noise of op-amps and kickback noise from comparators.

So the total output noise generated by sampling switches of SHA is:

\[
\overline{v_{in}^2} = \frac{2K_B T}{C_{s0} f_0}, \quad f_0 = \frac{C_{s0}}{C_{s0} + C_g}
\]  \hspace{1cm} (4.13)

The total output noise generated by sampling switches of MDAC is:

\[
\overline{v_{in}^2} = \frac{2K_B T}{(C_s + C_F) f_0 A_C}, \quad f = \frac{C_F}{C_s + C_g + C_F}
\]  \hspace{1cm} (4.14)

Thermal noise contribution of op-amps is:

\[
\beta = \sum_{\text{noise contributor}} g_m
\]

\[
\frac{4}{3} \frac{\beta K_B T}{[C_L + C_F (1 - f)] f_0 A_C^2}, f = \frac{C_F}{C_s + C_g + C_F}
\]  \hspace{1cm} (4.15)

So, the total output thermal noise generated by op-amps of SHA is:

\[
\overline{v_{in, \text{diff}}^2} = \frac{2K_B T}{C_{s0} f_0} + \frac{4}{3} \frac{\beta K_B T}{[C_{L0} + C_{s0} (1 - f)] f}
\]  \hspace{1cm} (4.16)

The total output thermal noise generated by op-amps of MDAC is:
\[
\overline{v_{in,k}^2} = \frac{2K_BT}{fA_C^2C_{F,k}} + \frac{4}{3}\beta K_BT \left[ \frac{C_{L,k}}{A_C} + C_{F,k}(1-f) \right] f
\]  
(4.17)

As the results, the total input referred thermal noise of ADC is:

\[
v_{\text{in,input}}^2 = v_{\text{in,SHA}}^2 + \sum_{k=1}^{M} \overline{v_{in,k}^2}
\]

\[
= \frac{2K_BT}{C_{0,0}} \left[ 1 + \frac{2\beta A_c}{3} \frac{1}{C_{0,0}} \right] + \sum_{k=1}^{M} \frac{2K_BT}{A_c^2} \left[ \frac{1}{C_{F,k}} + \frac{2\beta}{3} \frac{1}{C_{L,k} + C_{F,k}(1-f)} \right]
\]

\[
= \frac{2K_BT}{C_{0,0}} \left[ 1 + \frac{2\beta A_c}{3} \frac{1}{A_c} \right] \left[ \frac{1}{A_c-1} \right]^s + 4A_c(A_c-1) \alpha
\]

\[\sum_{k=1}^{M} \frac{2K_BT}{A_c^2} \left[ \frac{1}{A_c-1} + \frac{2\beta}{3} \frac{1}{A_c} \right] \left[ \frac{1}{A_c} \right]^{s^{k-1}} + \frac{1}{A_c} 4A_c(A_c-1) \alpha
\]

(4.18)

So all the input referred thermal noise must be smaller than quantizer error:

\[N_{\text{thermal}} = N_{\text{quant}} < \frac{\text{LSB}^2}{12}\]  
(4.19)

Then, we will get the minimum size of capacitors according to different scaling factors.

The next step is to find the optimized scaling factor \(s\). Let’s consume power consumption in SHA and MDACs is proportional to \(g_{m,\text{input}}\), then we can get the total \(g_m\) as follows:

\[
\tau = \frac{C_L + C_F(1-f)}{fB_m} \implies g_{m,k} = \frac{C_{L,k} + C_{F,k}(1-f)}{\tau_k f}
\]

\[
g_{m,\text{total}} = g_{m,0} + \sum_{k=1}^{M} g_{m,k} = \frac{C_{L,0}}{\tau_0} + \sum_{k=1}^{M} \frac{C_{L,k} + C_{F,k}(1-f)}{\tau_k f}
\]

\[
= \frac{C_s A_c}{\tau_0} \left[ \frac{s}{A_c-1} + 4(A_c-1) \alpha \right] + C_s A_c \left[ \sum_{k=1}^{M} \frac{A_c}{A_c-1} s^{k-1} + \frac{1}{A_c} s^k + 4A_c(A_c-1) \alpha \right]
\]

(4.20)
Figure 4.9 shows the power consumption with respect to capacitor scaling based on different stage partition. Figure 4.9(a) is the 2-bits per stage partition, and optimized scaling factor is 1/3; Figure 4.9(b) is the 3-bits per stage partition, and optimized scaling factor is 1/4; Figure 4.9(c) is the 4-bits per stage partition, and optimized scaling factor is 1/3; Figure 4.9(d) finds the overall optimized stage partition: 2-bits per stage is the best. Figure 4.9(e) shows the minimum capacitor of first stages with respect to scaling factor and different stage partition. As a result, the best choice of the most cases, if consider power optimization and thermal noise impact, is 2-bits per stage and the scaling factor should be 1/3.
Figure 4.9 Power vs. Capacitor Scaling Based on Different Stage Partition.
CHAPTER 5

PIPELINE ADC DESIGN EXAMPLE

5.1 State of art

There are various types of pipeline ADC design (sampling rate: 10M~500Msample/sec, resolution: 8~16 bits). Here only middle range sampling rate (around 40Msamples/sec) and middle range resolution (10 bits) are listed, since in the following section, one 40Msamples/sec, 10-bits pipeline ADC will present in order to verify the pipeline design methodology.
<table>
<thead>
<tr>
<th>Technology</th>
<th>Bits</th>
<th>Bits/Stage</th>
<th>SNDR [dB]</th>
<th>Speed [Msamples/sec]</th>
<th>Power [mA]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loloee 2002 [82]</td>
<td>12</td>
<td>1-1-1-1-1-1-1-1-1-1-2</td>
<td>65</td>
<td>80</td>
<td>144.44</td>
</tr>
<tr>
<td>Bogner [83]</td>
<td>14</td>
<td>3-3-2-2-4</td>
<td>64</td>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>Yoshioka [84]</td>
<td>10</td>
<td>1-1-1-1-1-1-1-1-1-1-3</td>
<td>56</td>
<td>80</td>
<td>13.3</td>
</tr>
<tr>
<td>Lee [85]</td>
<td>10</td>
<td>2-2-2-4</td>
<td>54</td>
<td>30</td>
<td>4.7</td>
</tr>
<tr>
<td>This work (3.3V IO)</td>
<td>10</td>
<td>1-1-1-1-1-1-1-1-1-1-2</td>
<td>55</td>
<td>40</td>
<td>12.7</td>
</tr>
</tbody>
</table>

Table 5.1 State of art
5.2 10-bits hybrid ADC specifications

Based on this pipeline ADC methodology study, this chapter presents a 10bit 40MHz power adaptive hybrid ADC with TSMC 90nmLP process: using op-amp sharing, dynamic biasing methods, it works in two modes: pipelined ADCs for high speed, cyclic ADC for low speed (only last stage runs, other stages are power off to save power). For pipeline mode, the total power consumption decrease as the sampling frequency drops.

This power adaptive Pipelined ADC has a target on applications in image sensor capturing, video processing, wireless communications, biomedicine, digital-intermediate frequency (IF) receivers and countless of digital devices. General pipelined ADC power consumption is almost the same no matter what sampling rate it works at. That’s because the total current of the most power consuming component, OP-AMP, maintains the same. It will waste lots of power if ADC sampling rate is not reach to maximum. This new pipelined ADC has self-adaptive switched-capacitor biasing circuitry, which will help ADC reduce power when it works at low sampling rate while maintaining the same performance.

The specification of this hybrid ADC is shown in table 5.2; and the pins definition of pipeline ADC is shown in table 5.3.
<table>
<thead>
<tr>
<th>Our specs</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>TSMC 90nm LP</td>
</tr>
<tr>
<td>Resolution</td>
<td>10 bits</td>
</tr>
<tr>
<td>Sampling Rate</td>
<td>maximum 40MHz/s</td>
</tr>
<tr>
<td>Power Supply</td>
<td>2.2V~3.0V For ADC core;</td>
</tr>
<tr>
<td>Input Range</td>
<td>1Vp-p Peak-to-peak input range</td>
</tr>
<tr>
<td>Power Dissipation</td>
<td>2~14mA Power listed excludes I/O buffer, reference voltage, and digital error correction power. Current is adaptive with sampling rate: For higher sampling rate, power consumption reaches to maximum.</td>
</tr>
<tr>
<td>DNL</td>
<td>Typical: +/- 0.3 LSB; Max: +/- 1 LSB Mismatching and offset error will be reduced by self-correction and digital calibration. So, the DNL and INL will be small enough to guarantee no missing code/level.</td>
</tr>
<tr>
<td>INL</td>
<td>Typical: +/- 0.5 LSB; Max: +/- 1 LSB</td>
</tr>
<tr>
<td>SFDR</td>
<td>53dB @ Low speed 50dB @ 40MHz/s</td>
</tr>
<tr>
<td>SNDR</td>
<td>58dB @ Low speed 53dB @ 40MHz/s</td>
</tr>
<tr>
<td>ENOB</td>
<td>9.3 bit</td>
</tr>
<tr>
<td>Output mode</td>
<td>CMOS</td>
</tr>
<tr>
<td>Digital Output</td>
<td>Offset binary output</td>
</tr>
</tbody>
</table>

Table 5.2 Hybrid Pipeline ADC Specification.
Mode_control signal controls the ADC function. When Mode_control=1, the ADC works in pipeline way; while, if the Mode_control=0, the ADC works in cyclic way.

<table>
<thead>
<tr>
<th>Pin</th>
<th>I/O type</th>
<th>Numbers</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1~D10</td>
<td>Output</td>
<td>10</td>
<td>10-bits digital output</td>
</tr>
<tr>
<td>VDD</td>
<td>Input</td>
<td>1</td>
<td>Power supply</td>
</tr>
<tr>
<td>GND</td>
<td>Input</td>
<td>1</td>
<td>Power supply</td>
</tr>
<tr>
<td>Mode_control</td>
<td>Input</td>
<td>1</td>
<td>ADC function controller</td>
</tr>
<tr>
<td>Vin+</td>
<td>Input</td>
<td>1</td>
<td>Differential signal input positive</td>
</tr>
<tr>
<td>Vin-</td>
<td>Input</td>
<td>1</td>
<td>Differential signal input negative</td>
</tr>
<tr>
<td>CLOCK</td>
<td>Input</td>
<td>1</td>
<td>Sampling clock</td>
</tr>
<tr>
<td>Vrefp</td>
<td>Input</td>
<td>1</td>
<td>Voltage reference</td>
</tr>
<tr>
<td>Vrefn</td>
<td>Input</td>
<td>1</td>
<td>Voltage reference</td>
</tr>
<tr>
<td>Vcom</td>
<td>Input</td>
<td>1</td>
<td>Voltage reference</td>
</tr>
<tr>
<td>Ibias</td>
<td>Input</td>
<td>1</td>
<td>Current biasing</td>
</tr>
</tbody>
</table>

Table 5.3 Pins Definition of Pipeline ADC.

5.3 Top level structure

Figure 5.1 shows the overall structure of hybrid ADC. It consist of 4stages MDAC (first 3 stages are identical, last stage works for pipeline and cyclic, its size is half
of the first 3 stages), dynamic biasing, control unit, non-overlapping clock generator and digital error correction.

![Figure 5.1 Hybrid ADC.](image)

5.3 **Sub-blocks structure**

This section talks all the important sub-block schematic of hybrid ADC, which includes clock generator, dynamic biasing, comparator, MDAC, op-amps.

5.3.1 **Non-overlapping clock generator**

The non-overlapping clock generator (Figure 5.2) is built by digital inverters and nand gates. It will generate time delay and make two sets of non-overlapping clock: ph1 and ph2. At the same time, it will also generate ph1e and ph2e, whose falling edges are a
little bit earlier than that of ph1 and ph2. Ph1e and ph2e will be used in MDAC for bottom sampling capacitance technology.

Figure 5.2 Non-overlapping Clock Generator.

5.3.2 Comparators

Section 3.2 shows three types of comparators. In this design, another comparator is used (shown in Figure 5.3). The advantage of this comparator is that it has self-biased voltage decision level:

Figure 5.3 Comparator.
\[ V_{ip} - V_{in} = \frac{(W/L)_1}{(W/L)_2} (V_{refp} - V_{refn}) \]  

(5.1)

5.3.3 Switched capacitor current source

Switched capacitor can work as a resistor; so, current could be modulated with sampling rate. At high sampling rate, op-amps in MDACs need more power to make accurate data acquisition; however, at low frequency, op-amps has more time to compute, as a result, its power could be reduced by adaptive biasing circuitry, at the same time, maintaining the same function.

The current of op-amps might vary 10 times if working at quite different sampling rate, which could have input differential pairs worked in weak saturation or triode region.
A control unit is applied to select whether ADC works pipelined normal mode or Low power cyclic mode.

5.3.4 OP-AMP

This design uses fold-cascoding op-amp, although telescopic op-amp is the most power efficient one. The reason is that fold-cascoding op-amp has larger output swing range. This is important, especially for dynamic biasing, because the current runs through transistors vary 10 times.
Figure 5.5 (a) Folded-cascade Op-amp with Gain Boosting (b) N-type Additional Op-amp (c) P-type Additional Op-amp.
5.3.4 MDAC

This design uses 1.5 bit per stage MDAC, and at the same time, using op-amp sharing. Traditional MDAC can only solve 1.5 bit per stage, and have 8 stages for 10 bit ADC. However, using op-amp, only 4 stages are enough. For traditional structures, op-amp only works at half the period: one half of the period for signal sampling to capacitors, at this time, op-amp does nothing; the other half of the period, op-amp amplifies the signals. So op-amp wastes half the power for waiting. Using op-amp sharing technology, one op-amp works for two stages, and it will work all the time. As shown in figure 5.6, now 1st, 2nd, 3rd stage are the same. They will get 6 sets of 2-bit output while only have 3 op-amps.

![Diagram of MDAC](image)

Figure 5.6 1st 2nd 3rd MDAC.
The last stage (Figure 5.7) works in pipeline and cyclic way. So, in pipeline mode, switches of ph2_1 and ph2_2345 are turn off. So no signal connects Vina_cyc/Vinb_cyc to op-amp. It works as normal pipeline ADC. However, if ADC works in cyclic way: these switches of ph2_1 and ph2_2345 are turn on according to scheme of figure 5.8.

Figure 5.7 4th MDAC.

Using 1~5 counter to control the switches during cyclic processing.

Figure 5.8 Phase Scheme of Cyclic ADC.
5.4 Overall layout

Figure 5.9 Layout.

103
As shown in figure 5.9, the overall area of hybrid ADC is 700um x 310um without pads.

## 5.5 Simulation results

Figure 5.10 shows the total power consumption with respect to sampling rate. If ADC works in pipeline mode, the total power consumption changes from 4.2 mA up to 12.2 mA, while the sampling rate increases from 4Msamples/second to 40 Msamples/second (blue curve). The black curve shows the total power consumption of cyclic mode, it’s much smaller than pipeline mode, and almost constant, since only 4\textsuperscript{th} stage are turn on, while other three stages and dynamic biasing circuitry are turn off.

![Figure 5.10 Total Power Consumption vs. Sampling Frequency.](image)

Figure 5.10 Total Power Consumption vs. Sampling Frequency.
If input signal is applied to a ramp signal, the output signal is upstairs signal. No-missing code occurs in Figure 5.11.

Figure 5.11 Ramp Signal to Verify Monotonicity.

Figure 5.12 shows the spectrum analysis of different input signal frequency. The input signal frequency is: 185546.875Hz, 400390.625Hz, 1005859.375Hz, 3994140.625Hz, 9970703.125Hz, and 19697265.63Hz, while the sampling rate is the maximum rate: 40Msamples/second. The number of samples is $2^{12} = 4096$ points. Using FFT analysis, the SNR, ENOB results as follows (Figure 5.13).
POWER SPECTRAL DENSITY PLOT

SNR = 61.56 dB @ OSR = 16

(a)

POWER SPECTRAL DENSITY PLOT

SNR = 61.44 dB @ OSR = 16

(b)
Figure 5.12 Spectrum Analysis.
Figure 5.13 SNR and ENOB.
CHAPTER 6

CONCLUSIONS

6.1 Conclusions

The thesis presents a comprehensive pipeline ADC design methodology. The pipeline ADC design methodology provides a comprehensive and quantitative mapping matrix between system level ADC performance specs (e.g., power optimization, sampling rate, resolution, input voltage range, etc.) and the critical design parameters at block levels, such as, thermal noise limitation, op-amp selection in S/H and MDAC, DC gain and closed-loop gain bandwidth of op-amps requirement, sampling capacitor selection criterion etc.

The design methodology was verified by designing a 10-bit 40Msample/second hybrid ADC using TSMC 90nm IO device process. This hybrid ADC uses op-amp sharing method to save half number of stages, and, it can work in two modes: pipelined ADCs for high speed, cyclic ADC for low speed (only last stage runs, other stages are power off to save power). It also adapts dynamic biasing methods, so, for pipeline mode, the total power consumption decrease as the sampling frequency drops.
6.2 Future Work

First, offset of op-amp worse the pipeline ADC performance. For traditional MDAC, it’s easy to implement offset cancellation technology, such as auto-zeroing, since op-amps works only in half of the period. So it can store the offset at the other half of the period, and subtract the offset then. While, using op-amp sharing method, op-amps are always working. So auto-zeroing cannot implement into it. Some advanced technology should be considered, such as chopping.

Second, digital calibration [38, 39, 42, 54, 56, 62, 65, 72, 78, 79, 81] would help improve ADC performance: SNR, SFDR, mismatching and so on. Based on this hybrid ADC design, new digital calibration scheme should be developed in the future in order to improve ADC performance.
REFERENCE


[54] Seung-Chul Lee; Kwi-Dong Kim; Jong-Kee Kwon; Jongdae Kim; Seung-Hoon Lee, "A 10-bit 400-MS/s 160-mW 0.13um CMOS Dual-Channel Pipeline ADC Without


[82] Arash Loloee, Alfio Zanchi, Huawen Jin, Shereef Shehata, Eduardo Bartolome, "A 12b 80MSps Pipelined ADC Core with 190mW Consumption from 3V in 0.18µm Digital CMOS", ESSCIRC 2002.


[85] Seung-Chul Lee, Kwi-Dong Kim, Jong-Kee Kwon, Jongdae Kim, "A 4.7mW 0.32mm² 10b 30MS/s Pipelined ADC Without a Front-End S/H in 90nm CMOS", ISSCC 2007.