Design of Multi-GHz Data Converter Components

**A dissertation submitted in partial fulfillment of the
requirements for the degree of
DOKTORINGENIEUR (Dr.-Ing.)
in
Electrical Engineering**

**to
Faculty of Computer Science, Electrical Engineering and
Mathematics**

University of Paderborn

**by
Master of Science Samiran Halder
Kolkata, India**

Design of Multi-GHz Data Converter Components

Zur Erlangung des akademischen Grades
DOKTORINGENIEUR (Dr.-Ing.)
der Fakultät für Elektrotechnik, Informatik und Mathematik
der Universität Paderborn
vorgelegte Dissertation
von
Master of Science Samiran Halder
aus Kalkutta, Indien

Design of Multi-GHz Data Converter Components

**Zur Erlangung des akademischen Grades
DOKTORINGENIEUR (Dr.-Ing.)
der Fakultät für Elektrotechnik, Informatik und Mathematik
der Universität Paderborn
vorgelegte Dissertation
von
Master of Science Samiran Halder**
aus Kalkutta, Indien

**Referent: Prof. Dr.-Ing. Andreas Thiede
Korreferent: Prof. Dr.-Ing. Rolf Kraemer**

Tag der mündlichen Prüfung: ....

Paderborn, den

Diss. EIM-E/253

DECLARATION

This is to declare that the thesis entitled **"Design of Multi-GHz Data Converter Components"** submitted by Mr. Samiran Halder to the "Faculty of Computer Science, Electrical Engineering and Mathematics" of "University of Paderborn" for the partial fulfillment of the award of **"DOKTORINGENIEUR (Dr.-Ing.)".** Mr. Samiran Halder was awarded Master of Science degree from Indian Institute of Technology, Kharagpur, India, in 2005. This thesis was prepared by Mr. Samiran Halder under the supervision of Prof. Andreas Thiede of the "Faculty of Computer Science, Electrical Engineering and Mathematics" of "University of Paderborn". The results embodied in this thesis have not been submitted for any other degree or diploma in any other university or institution.

Date:

(Samiran Halder)

Dedicated to my parents

Abstract

Design of Multi-GHz Data Converter Components
by
Samiran Halder

In the last few decades the communication bandwidth has evolved with an enormous speed and the requirement for high-speed data converters is directly dictated by that. In RF systems, the analog-digital interface is pushed towards the antenna, because the complex signal processing can be handled more efficiently in the digital domain. On the other hand it makes the design of these high-speed data converters more and more difficult.

In this dissertation the main design challenges in the field of multi-GHz data converters are discussed. The main research work is broadly divided into two parts. In the first part the different design techniques of multi-GHz analog to digital converters (ADC) are presented. In the other section the design of multi-GHz current steering digital to analog converters (DACs) have been discussed.

In the context of ADC design the front-end track and hold (THA) comes as the most critical part. This is because of the fact that any error introduced in this block cannot be compensated by the signal post-processing. In this research work an attempt has been made to improve the performance of the THA so that the stringent accuracy requirements of the quantization process can be relaxed. This is accomplished by enhancing the input range of the THA. Two different kind of THAs are developed. In both the THAs, different techniques are used to enhance the input range up to 2Vpp differential at the sampling rate of 10GHz. According to the authors knowledge these THAs are the only published THAs which can work with 2Vpp input signal and achieve an accuracy of more than 6.5-bit at a sampling rate of 10GHz. A new double sampled technique is proposed for the open loop THA architectures which can be instrumental to double the sampling speed of the THA with a little overhead of power dissipation compared to conventional open loop THAs. As a design example a 20GHz 6-bit comparator has been designed and measured successfully.

An 8-bit segmented current steering DAC has already been designed. As a tread-off between the accuracy and power consumption 50% segmentation is used. The MSB sub-DAC is implemented with conventional unary weighted DAC architecture. In the context of high-speed DAC design the binary to thermometer decoder comes as the design bottleneck in terms

of speed and power. In this unary sub-DAC design a novel thermometer decoder is proposed which is mainly based on an HBT ROM structure. In simulation the 8-bit DAC shows an accuracy of 7.83 effective number of bits (ENOB) with 9GHz of single tone input sinusoidal and a sampling rate of 20GHz. The 4-bit LSB sub-DAC is already implemented with a weighted resistive ladder network. A novel binary weighted resistive ladder network is proposed. The 4-bit DAC is found to be functional up to 30GHz of sampling rate which shows the second best performance in terms of sampling speed for published SiGe high-speed DACs.

ACKNOWLEDGEMENTS

I will always consider my experience of the past three and half years in the department of Circuit design, IHP Microelectronics GmbH, as one of the most wonderful and enjoyable parts of my life. I am so deeply indebted to many people during the work leading to this dissertation. First of all, I would like to express my sincere gratitude to Prof. Rolf Kraemer and, Prof. Andreas Thiede who has encouraged and guided me through all my research work. They are true advisor to me. I will benefit from their precious and generous advices lifelong.

I am grateful to Dr. Hans Gustat who always leave his doors open for me whenever I encountered technical problems or need his suggestion on my work. During my research work in IHP Microelectronics GmbH, he was my project leader and it would not be possible to accomplish the dissertation without his constant encouragement. He is virtually a mentor to me.

I owe my gratitude to Dr. Christoph Scheytt, head of the department Circuit design, for providing the opportunity to pursue my research work in good friendly atmosphere. His technical advices and insights are really valuable to the success of this research.

I would say thanks to all of the team members of department of Circuit design, which include Yaoming Sun, Sabbir A. Osmany, Kai Hu, Dr. Frank Herzel, Dr. Wolfgang Winkler, Dr. Klaus Schmalz. I am so fortunate to be able to work with such a group of extraordinary colleagues and good friends. I would like to specially thank Mr. Yevgen Borokhovych whose talent and persistence assured our achievements in the high-speed project.

The last but not the least people to mention are the most important ones in my life, my family. The accomplishment of this work would not be possible without them. My parents and sister are the great source of support and encouragement to my research and life. I cannot adequately express the love and gratitude I feel for them.

Index

List of Figures

List of Tables

Chapter 1
Introduction

1.1. Motivation

Wireless communications have been the driving force in analog electronics development during the last decades. As the end products are produced for every-day use, the price, size, and weight of the devices play a large part in determining their design. Cost reduction and miniaturization require higher integration levels. Reasons for a high level of integration are increased reliability and product security.

In Fig.1.1 the requirements of data converters for different applications are plotted. Very high accuracy data converters are used in the slow instrumentation purposes, whereas the main applications of data converters are dominated by the communication systems. Wireless communication standards, such as Universal Mobile Telecommunication System (UMTS), Wireless Local Area Network (WLAN), Wireless Local Loop (WLL) or Local Multipoint Distribution Services (LMDS), are evolving towards higher data rates, thus allowing more services to be provided. In addition to that in communication radios, the analog digital interfaces are pushed towards the antenna as signal processing can more conveniently be done in digital domain. Currently almost all data converter vendors are providing analog to digital converter (ADC) or digital to analog converter (DAC) based on CMOS solutions with sampling rates at 1GHz and above. In some upcoming applications e.g. satellite or radar communication systems basestation applications with low to medium resolution data converters and multi-GHz sampling rate are going to be used. Such data converters are also very useful for broadband measuring instruments such as sampling oscilloscopes or arbitrary signal generators.

The state-of-the-art CMOS technology falls short to meet such stringent requirement for the data converters, where the SiGe BiCMOS technology quintessentially has an edge. While high-speed requirements can be fulfilled with the faster HBT devices, low-speed blocks can be implemented with CMOS devices. In this work the design aspects of high-speed data converts in 0.25µm SiGe BiCMOS technology are discussed. The main goals of this work are firstly, to investigate the dominating error sources in the data converter which restrict their performance, secondly to develop techniques to cope with those problems. Finally, to come

up with a set of data converter components which can be used as standalone systems as well as the building blocks for the complex ADC or DAC.
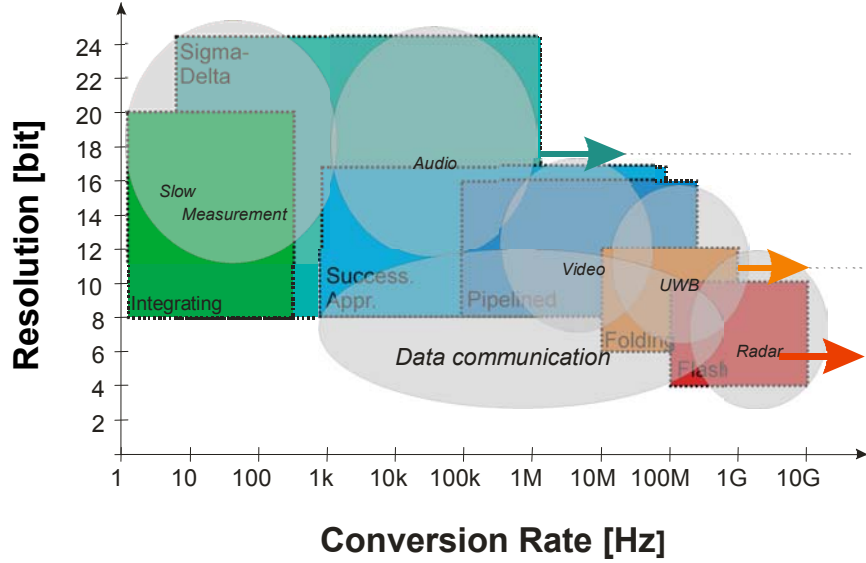
Fig. 1.1. Requirements of data converters for different applications

1.2. Research Contribution

In section 1.1 it is discussed that the upcoming applications require multi-GHz of sampling rate data converters with low to medium resolution. In the thesis, an attempt has been made to develop some of the key components of ADC and DAC which can be used as a standalone system and as sub-blocks to build up complex high-speed, high accuracy data converters.

In the context of ADC design, the front-end track and hold amplifier (THA) comes as the bottleneck for the full system. This is because of the fact that any error introduced in this stage cannot be corrected by post processing the sampled analog signal. In this work an attempt has been made to improve the performances of the THA so that the tough requirements of the quantizer block can be relaxed. Two different kinds of THAs are implemented and measured successfully. In both the THAs, different techniques are used to enhance the input range up to 2Vpp differential at the sampling rate of 10GHz. To accomplish this requirement, the input buffer of the THAs are optimized. For the first time a cascode input buffer is used in the open loop THA design,while in another variant of THA a new complementary npn and pnp emitter follower is used. According to the authors knowledge these THAs are the only published THAs which can work with 2Vpp input signal and achieve an accuracy of more than 6.5-bit at a sampling rate of 10GHz.

2

A new double sampled technique is proposed for the open loop THA architectures which can be instrumental to double the sampling speed of the THA with a little overhead of power dissipation compared to conventional open loop THAs. A novel double sampling switch is proposed which will make the sampling process insensitive to the clock skew, that appears as the bottleneck for the double sampling THAs and restricts the resolution.

With the advent of modern wireless communication systems different direct signal synthesis techniques are emerging as very popular. In this kind of systems the front-end DAC comes as the crucial component. The DAC should have low power dissipation and the resolution ranges from 4-12-bits. In this work an attempt has been made to design current steering DACs with a resolution of 4-8-bits and a sampling frequency ranging from 20GHz-30GHz. The 4-bit DAC is implemented with a weighted resistive ladder network. A novel binary weighted resistive ladder network is proposed. The 4-bit DAC is found to be functional up to 30GHz which shows the second best performance in terms of sampling speed for published SiGe high-speed DACs.

An 8-bit segmented current steering DAC has already been designed, where the 4-bit 30GHz DAC is used as the LSB sub-DAC. The MSB sub-DAC is implemented with conventional unary weighted DAC architecture. In the context of high-speed DAC design the binary to thermometer decoder comes as the design bottleneck in terms of speed and power. In this unary sub-DAC design a novel thermometer decoder is proposed which is mainly based on an HBT ROM structure. In simulation the 8-bit DAC shows an accuracy of 7.83 effective number of bits (ENOB) with 9GHz of single tone input sinusoidal and a sampling rate of 20GHz.

1.3. **Organization of Thesis**

In chapter 2 a brief review on different ADC architectures particularly suitable for high-speed applications are presented. Different static and dynamic parameters for ADCs are defined. A general trend of ADC performance improvement with time is discussed. Finally the advantages and disadvantages of different ADC architectures are critically analyzed.

In chapter 3 the design techniques for high-speed open loop THAs are presented. After a brief review of the most commonly used THA architectures the design methods of two different THAs with optimized high input swing are described. A new open loop double sampling THA architecture is presented to enhance the sampling speed of the THA. Finally the tested results of the implemented THAs are presented.

The design of a high-speed comparator is also described in chapter 3. Different common error sources of open loop comparators are analyzed. The design of a conventional open loop comparator is presented. Finally the measurement results of a 20GHz comparator are presented.

An architecture review of current steering DACs is presented in chapter 4. Different error sources associated with the current steering DAC are analyzed. The state-of-the-art techniques to enhance the static and dynamic performances are presented. But these techniques are found to be not very useful for the multi-GHz DAC design. Thus, a novel non-binary weighted DAC based current cell calibration technique is proposed which can be useful to enhance the performance of high-speed DACs.

The designs of a 4-bit as well as an 8-bit current steering DAC have been presented in chapter 5. In the 4-bit DAC implementation a novel resistive weighting network in used. The design of a 20GHz 8-Bit segmented current steering DAC is presented afterwards. A new HBT ROM based thermometer decoder architecture is proposed which could be instrumental to enhance the speed and latency requirements for high-speed unary weighted DACs . For the 8-bit DAC simulation results are presented. Finally measurement results of the 4-bit DAC and a brief comparison with the state-of-the-art multi-GHz SiGe DAC are presented.

Finally the conclusions are drawn in chapter 6. This chapter also includes the future scope of the research work.

Chapter 2
Analog to Digital Converter

_____

2.1. Introduction

The requirement of high-speed, high-resolution analog to digital converters (ADC) is directly dictated by the evolution of modern communication systems. Ultra-wideband and radar communication systems are going to use ADCs with a sampling rate of few gigahertz to few tens of gigahertz. Designing such high-speed ADCs with moderate resolution becomes a great challenge for silicon-germanium (SiGe) technology.

The basic analog to digital conversion can be considered as summation of two main operations (see Fig. 2.1). The first operation is called sampling. In this process the continuous time analog signal is converted into a discrete time analog signal. After this, the sampled analog signal is approximated to some predefined discrete amplitudes. This process is known as quantization. Each of the discrete analog amplitudes is then assigned to a specific digital code.
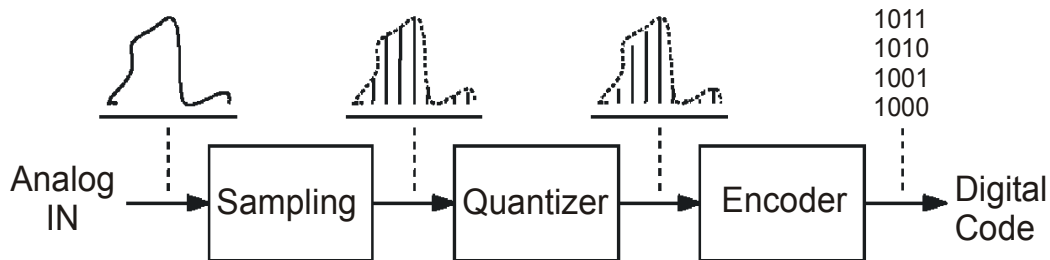
Fig. 2.1. Analog to digital conversion

In this chapter the basic quantization process is discussed in section 2.2. The static and dynamic errors of the analog to digital converter are defined in section 2.3. A present scenario and design trends of ADC design are presented in section 2.4. The architectures of different ADCs which can be used for gigahertz range sampling rate are presented in section 2.5. Finally the conclusions are drawn in section 2.6.

2.2. Quantization

The quantization process can be defined as mapping of time discrete analog signal into a finite set of digital words. As mentioned in the earlier section the basic A/D conversion process can be characterized as sampling the continuous signal in the time domain and then assigning the time discrete amplitudes into some digital code words, i.e. quantization. In spite of this sometimes the terms A/D conversion and quantization are used synonymously. A quantizer can be uniquely described by its transfer function or quantization characteristic, which indicates the discrete outputs as a function of the continuous input signal. The quantization characteristic therefore contains two sets of information: the first includes the digital codes associated with each output state, and the second includes the threshold levels which are the set of input amplitudes at which the quantizer transits from one output code to the next (Fig. 2.2). Various kind of digital coding can be used. Those are namely natural binary, sign plus magnitude, offset binary, one's complement, two's complement, binary coded decimal (BCD), and Gray code; each of the coding scheme has its own advantages in particular application.

In a quantizer, if there exists M threshold levels, that will generate (M+1) output digital code words. The threshold levels are denoted by T_k, where k ranges from 1 to M. Quantization step (Q) is defined as,

$$Q = T_{k+1} - T_k \qquad (2.1)$$
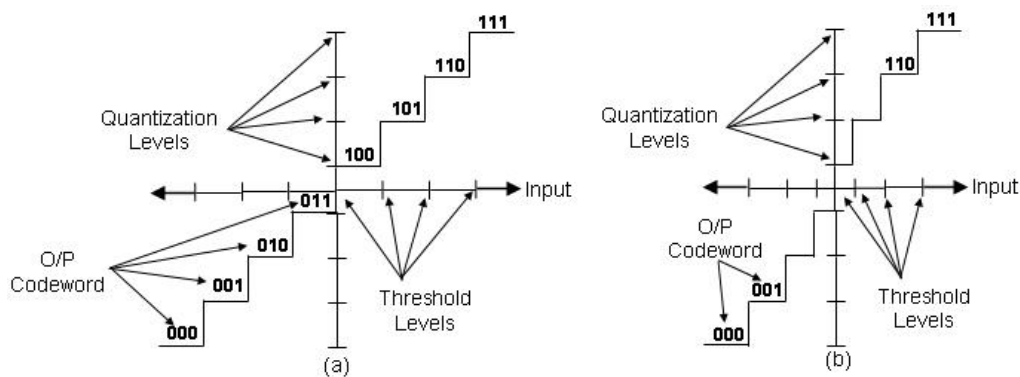
Fig. 2.2 Transfer characteristics of (a) uniform, (b) nonuniform quantization

The ideal threshold levels are denoted by T_k^*. This ideal threshold levels can be spread over the abscissa of the quantizer transfer function. As shown in Fig. 2.2(a) the quantization steps are equal. This kind of quantization is known as uniform quantization otherwise it is termed as nonuniform quantization as shown in Fig. 2.2(b). The optimum performance results

in when the threshold locations matche with the probability distribution function of the incoming signal. However, in the absence of a priori knowledge of the input signal statistics, uniform quantization outperforms other arrangements. Therefore, uniform quantizers are most commonly used.

Depending upon the location of origin the quantization process can be classified into two categories. In bipolar quantization the ideal threshold levels are spread symmetrically about the origin (Fig. 2.3a). On the contrary in unipolar quantization the threshold levels are placed either positive or negative direction with respect to the origin. In Fig. 2.3b an example of unipolar quantization presented.

The Full-Scale Range, FSR, of a uniform quantizer represents that portion of the transfer function domain spanned by all equal length intervals (M) between adjacent ideal thresholds. Thus the quantization step (Q) can be alternatively defined as,

$$Q = \frac{FSR}{M}$$
(2.2)

(a)   (b)
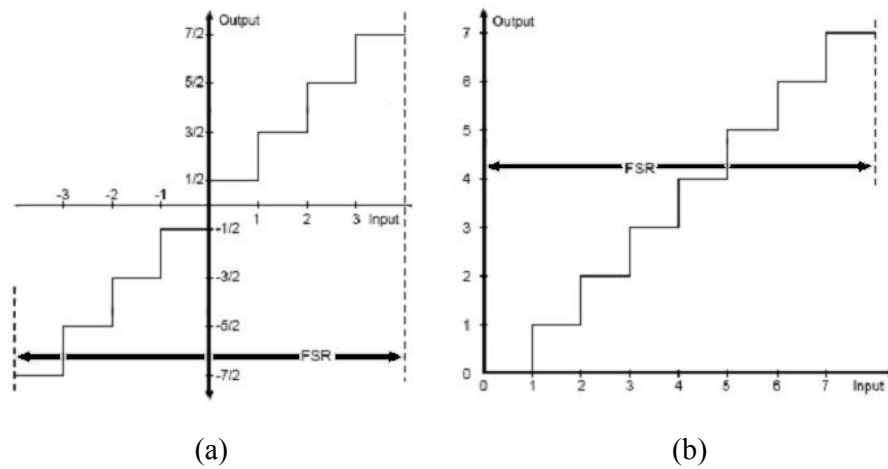
Fig. 2.3. Transfer function of (a) bipolar (b) unipolar quantizatioin

In Fig. 2.4 two of the most commonly used quantization transfer characteristics are presented. Those are known as mid-tread and mid-riser characteristics. For an N-bit bipolar or unipolar quantizer, mid-tread quantizer has $M=2^N-1$ no. of quantization levels with a quantization level at origin in case of bipolar quantization or at FSR/2 for unipolar quantization. Mid-riser characteristics $M= 2^N$ no. of quantization levels with a threshold value at the origin for bipolar quantization (FSR/2 for unipolar quantization.). Thus practically for an N bit quantizer (M-1) threshold levels are required. In mid-riser characteristics $M=2^N$ quantization levels directly mapped into 2^N binary codes. For this reason the mid-riser

quantizer is more popular than the mid-read counterpart. In fig. 2.4b an ideal mid-riser transfer characteristic is shown for 3-bit quantizer. The quantization step Q for mid-riser quantizer is given by,

$$Q = \frac{FSR}{2^N} \tag{2.3}$$
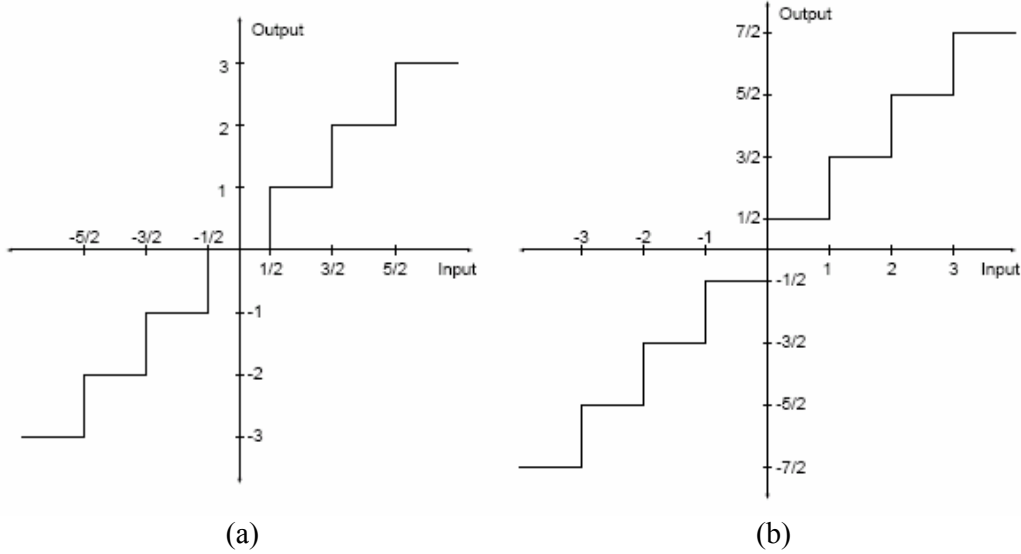
Fig. 2.4. (a) Mid-tread, (b) Mid-riser quantizer

2.3.1. Static Errors in ADC

Due to the imperfection in fabrication the real quantization transfer function deviates from the ideal one. The actual thresholds (T_k) have some error with respect to their ideal placements (T_k^*). Such non-idealities are known as static or DC errors and can be defined in several ways. The definitions of static errors of quantizer are indicated in the transfer curve of a converter. As shown Fig. 2.5(a) the error which causes an equal amount of shift in all the thresholds is known as offset error of the quantizer. Non-ideality which causes a same amount of step size error for all of the quantization steps is known as gain error. The gain error is shown in Fig. 2. 5b.

The most important measures of static error of quantizers are indicated by integral nonlinearity (INL) and differential nonlinearity (DNL). These properties actually indicate the accuracy of a converter and include the errors of quantization, nonlinearities, short-term drift offset and noise. Integral nonlinearity (INL), sometimes called relative accuracy, is defined as

the deviation of the output code of a converter from its ideal counterpart excluding a possible offset error. The nonlinearity should not deviate more than ±1/2 LSB from the ideal transfer curve. This INL boundary implies a monotonic behavior of the converter. Monotonicity of an analog-to-digital converter means that no missing codes can occur [1].
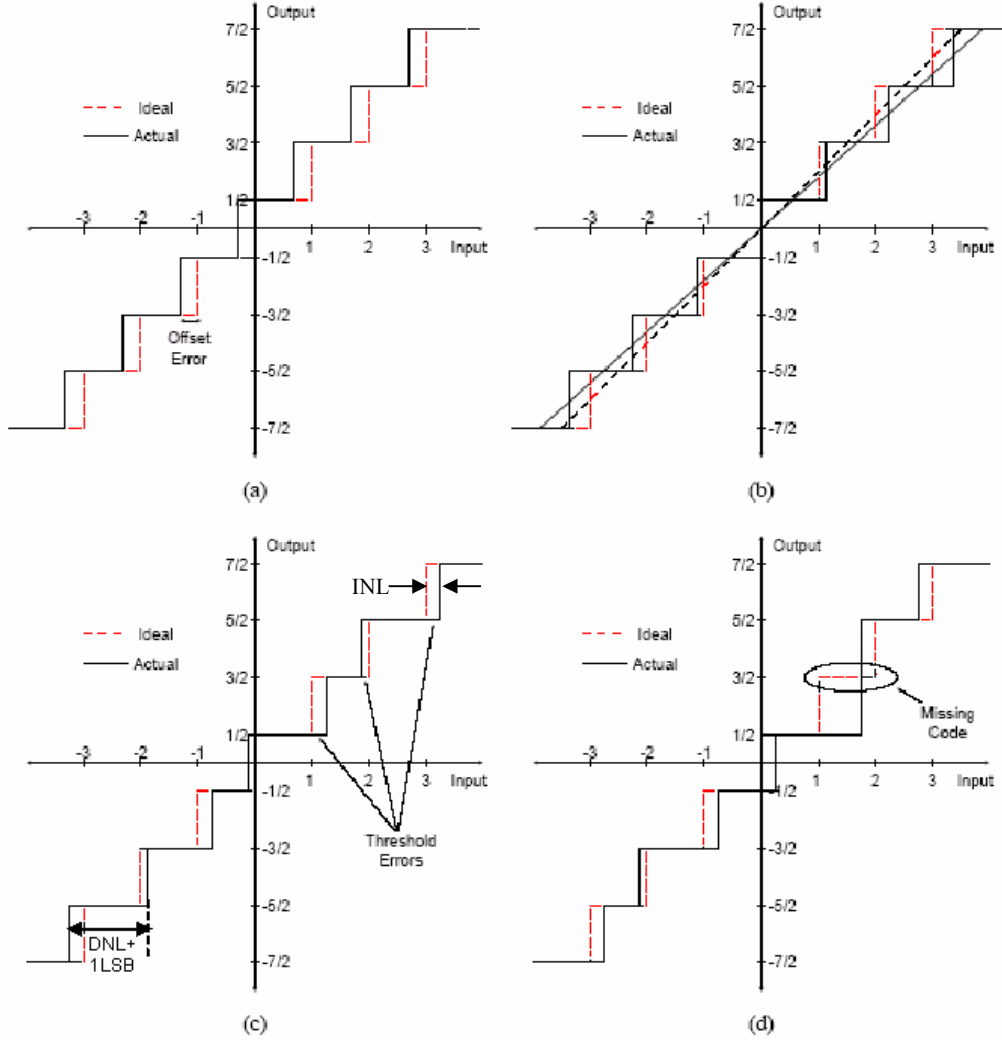
Fig. 2.5 (a) Offset error, (b) Gain error, (c) Threshold errors (INL & DNL), (d) Missing codes

Differential nonlinearity (DNL) error gives the difference between two adjacent threshold values (T_k, T_{k-1}) compared to the quantization step (Q) of a converter generated by transitions between adjacent pairs of digital code numbers (D_k) over the whole range of the converter. The DNL of ADC output D_k can be written in terms of LSB as,

$$DNL(D_k) = \frac{T_k - T_{k-1} - Q}{Q} \tag{2.4}$$

There is a direct connection between the INL and DNL. The INL for output code D_k can be obtained by summing the DNL until code k,

9

$$INL(D_k) = \sum_{i=1}^{k} DNL(D_k) \qquad (2.5)$$

2.3.2. Dynamic Errors in ADC

Dynamic performance parameters include information about noise, dynamic linearity, distortion, settling time errors, and sampling time uncertainty of an ADC. It should be noted that all the measures following are both frequency and signal amplitude dependent. Furthermore, unless otherwise specified, they are obtained with a full-scale input signal.

2.3.2.1. Signal-to-Noise ratio (SNR)

The quantization process introduces an irreversible error, which sets the limit for the dynamic range of an A/D converter. Assuming that the quantization error of an ADC is evenly distributed over any quantization level, the power of the generated noise in a 1Ω is [1] given by,

$$e^2 = \frac{Q^2}{12} \qquad (2.6)$$

where, e^2 is the quantization noise power and Q is the quantization step. If a single-tone sine wave signal with maximum amplitude is adopted for a full scale range (FSR) of a quantizer with a large number of bits (N≥5), the signal power is given by,

$$S_p = \frac{FSR^2}{8} \qquad (2.7)$$

Combining Eq. 2.3 and Eq. 2.7 and substituting in Eq. 2.6 the signal-to-noise ratio (SNR) for a single-tone sinusoidal signal can be obtained to be,

$$SNR = 2^{2N} \bullet \frac{3}{2} \qquad (2.8)$$

SNR can be expressed in dB by the following equation,

$$SNR = (6.02N + 1.76) \, \text{dB} \qquad (2.9)$$

When determining the SNR, the ratio between the frequency of the sine wave and the sampling frequency should be irrational. If the input signal deviates from the sine wave, the constant term, which depends on the amplitude RMS value of the waveform, differs from 1.76 dB. Eq. 2.9 indicates that each additional bit, N, gives an enhancement of 6.02 dB to the SNR. If oversampling is used, which means that the sample rate f_s is much larger than the

10

signal bandwidth f_{sig}, the quantization noise is averaged over a larger bandwidth and the signal-to-noise ratio becomes larger, written as,

$$SNR = 2^N \sqrt{\frac{3}{2}} . \sqrt{OSR} = (6.02N + 1.76 + 10\,log(OSR))\,\text{dB} \qquad (2.10)$$

where, the oversampling ratio is given by OSR,

$$ORS = \frac{f_s}{2 \bullet f_{sig}} \qquad (2.11)$$

In the Nyquist rate A/D converters, the signal bandwidth is normally equal to $f_s = 2 \bullet f_{sig}$ resulting in an OSR equal to one, while Eq. 2.10 suggests that the signal-to-noise ratio increases by 3 dB per octave of oversampling.

2.3.2.2. Total Harmonic Distortion (THD)

Any nonlinearity in an ADC creates harmonic distortion. In differential implementations, the even order distortion components are ideally canceled. However, the cancellation is not perfect if any mismatch or asymmetry is present. The total harmonic distortion (THD) describes the degradation of the signal-to-distortion ratio caused by the harmonic distortion. By definition, it can be expressed as an absolute value with,

$$THD = \frac{\sqrt{\sum_{j=2}^{(H+1)} V^2(j \bullet f_{sig})}}{V(f_{sig})} \qquad (2.12)$$

where, H is no. of harmonics to be considered and $V(f_{sig})$, $V(j \bullet f_{sig})$ are the amplitudes of fundamental and jth harmonic respectively.

2.3.2.3. Signal to Noise and Distortion Ratio (SNDR)

A more realistic figure of merit for an ADC is the signal-to-noise and distortion ratio (SNDR), which is the ratio of the signal energy to the total error energy including all spurs and harmonics. SNDR is determined by employing the sine-fit test, in which a sinusoidal signal is fitted to a measured data and the errors between the ideal and real signal are integrated to get the total power of noise and distortion [2],[3]. If all tones and spurs other than the harmonic distortion are considered as noise, the signal-to-noise ratio can be obtained from the SNDR by subtracting the total harmonic distortion from it

$$SNR_{real} = SNDR - THD \qquad (2.13)$$

where SNDR and THD are given in absolute values.

2.3.2.4. Spurious Free Dynamic Range (SFDR)

In wireless telecommunication applications, large oversampling ratios are often used and the spectral purity of the A/D converter is important. For such situations, a proper specification is the ratio between the powers of the signal component and the largest spurious component within a certain frequency band, called spurious free dynamic range (SFDR). The SFDR is usually expressed in dBc as,

$$SFDR(dBc) = 10.log\left(\frac{V^2(f_{sig})}{V^2(f_{spur})}\right) \qquad (2.14)$$

where $V(f_{sig})$ is the amplitude of the fundamental sinusoidal input and $V(f_{spur})$ the amplitude of the largest spurious. For an exact SFDR definition, the power level of the fundamental signal relative to the full-scale must also be given. Normally the limiting factor of the SFDR in ADCs is harmonic distortion. In most situations, the SFDR should be larger than the signal-to-noise ratio of the converter [4].

2.3.2.5. Effective Number of Bits (ENOB)

In ideal ADCs, the maximum analog bandwidth is equal to half the sampling bandwidth, according to the Nyquist theorem. The effective resolution bandwidth (ERB) is defined as the maximum analog frequency for which the signal-to-noise ratio of the system is decreased by 3 dB or 1/2 LSB with respect to the theoretical value. For a single-tone full-scale sinusoidal test signal with the maximum frequency within the ERB the effective number of bits (ENOB) can be defined according to the following equation,

$$ENOB = \frac{SNDR(dB) - 1.76}{6.02} \qquad (2.15)$$

where, SNDR is taken as the figure of merit to calculate the ENOB. Although according to the requirements of different applications the ENOB can also be estimated by considering either of SNR, THD or SFDR as the measure of linearity.

2.3.2.6. Dynamic Range

Dynamic range (DR) is the input power range for which the signal-to-noise ratio of the ADC is greater than 0 dB. The dynamic range can be obtained by measuring the SNR as a function of the input power.

2.4.   Performance Analysis and Present Trends in ADC Design

In the previous section the definitions of different static and dynamic parameters of ADC are presented. In this section an attempt has been made to analyze the ADC performances according those parameters. Although a comprehenive performance can be presented by the following set of parameters: stated resolution, sampling rate, SNR, SFDR and the power dissipation.

The pioneering work regarding the facts and trends of ADCs in terms of different dynamic performances was presented in [5] by R.H. Walden. The work has two main aspects; different physical processes which defines the upper or lower limits of ADC and the performance improvement in terms of time of ADC development. It revels a very interesting relation between the sampling rate and ENOB: resolution is decreased by 1-bit with doubling the sampling rate.

To analyze the performance of the ADC, SNR bits are taken as a figure of merit. The SNR bits are defined as follows,

$$SNR_{bits} = \frac{SNR(dB) - 1.76}{6.02} \tag{2.16}$$

It is been observed that the difference between the stated resolution of ADC and the SNR bits is about 1.5-bits. This is attributed to the nonlinearity and noise sources associated with different components of the ADC.

In the Fig. 2.6 the performance of different published and commercially available ADCs in terms of sampling frequency and the SNR bits are presented with the updated entries up to the year of 2005. Along with that, it also depicts the different physical horizons which dictate the limits for the ADC dynamic performances. The main parameters which influence the dynamic performances are thermal noise floor, aperture uncertainty in the sampling process and the comparator ambiguity or the comparator metastability. The last two effects will be discussed in details in chapter 3. The relation between these error sources and the maximum achievable SNR bits are derived in [5].

13
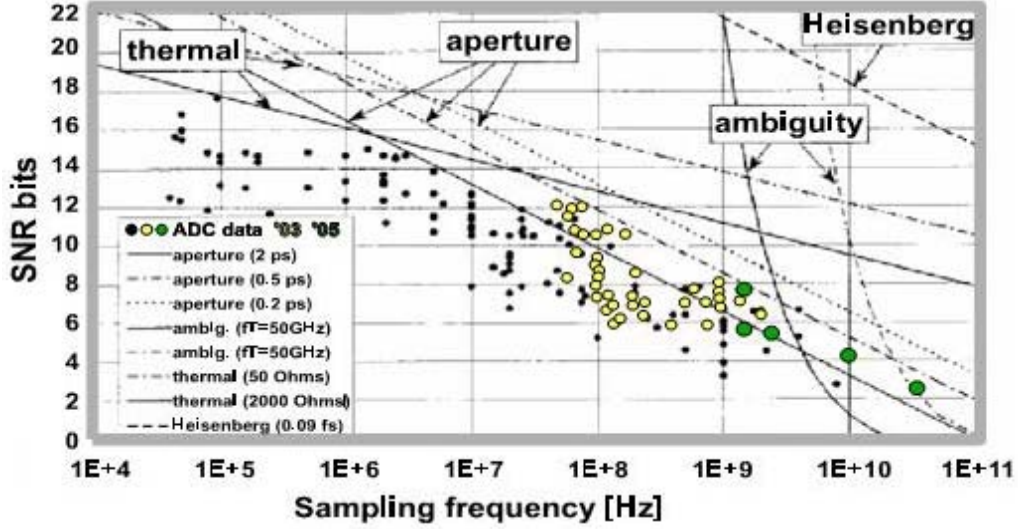
Fig. 2.6. Performance limits of ADC due to different physical phenomena [5]

The relation between the input referred thermal noise and the maximum attainable resolution in terms of SNR bits can be expressed as follows,

$$N_{thermal} = \log_2 \left(\frac{V_{FS}^2}{6kTR_{eff} f_{sample}} \right)^{\frac{1}{2}} - 1 \tag{2.17}$$

Where the $N_{thermal}$ represents the maximum SNR bits which can be achieved for a given equivalent input referred noise resistance (R_{eff}). V_{FS} is the full scale voltage of the ADC and f_{sample} is the sampling rate. T = Temperature in Kelvin. Boltzmann's constant, $k = 1.38658 * 10^{-23}$ J/K.

Assuming that the rms aperture uncertainty τ_a is known. Then the upper limit of the SNR bits ($N_{aperture}$) for the given τ_a can be presented as follows,

$$N_{aperture} = \log_2 \left(\frac{2}{\sqrt{3}\pi f_{sample} \tau_a} \right) - 1 \tag{2.18}$$

The relation between the SNR bits and the comparator ambiguity can expressed as,

$$N_{ambiguity} = \frac{\pi f_T}{6.93 f_{sample}} - 1 \tag{2.19}$$

Eq. 2.19 relates the transit frequency (f_T) of devices in a particular technology with the maximum resolution that can be achieved ($B_{ambiguity}$) for the sampling rate of f_{sample}.

The ultimate limit of the ADC resolution and sampling rate is estimated from Heisenberg uncertainty principle. This defines the least amount of resolvable energy corresponding to the ½ of LSB which can be detected in a given time interval i.e. the half of the sampling period.

This limit is almost four orders of magnitude higher than the state of the art ADC reported as of now.

The performance envelop of ADC is shifting but in a much lower rate than the technology evolved. It is only 1.5 bits in every 8 years as indicated in [5]. Although in [5] ADCs up to 1997 have been considered, but the present scenario remains almost same. The main reason behind that could be the traditional approaches to solve the analog problems like device mismatch in the analog domain. Digital post processing may be helpful for this sort of problem but increases the complexity of the full system.
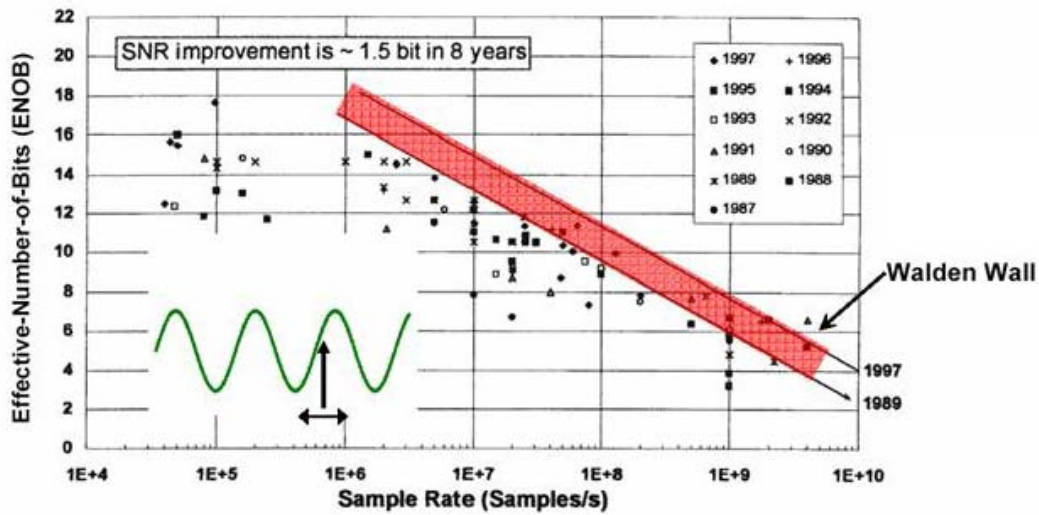
Fig. 2.7. Performance envelop improvement of ADC [5]

The performances of different ADC architectures are presented in [6]. It shows the highest resolution is achieved in sigma delta architectures. But it uses the large amount of oversamplng ratio. The optimum performance in terms of sampling rate, resolution and power can be obtained by using pipeline architecture. But this is very difficult to implement for gigahertz range sampling regime. Flash architecture is the fastest and comes as the obvious choice for higher sampling rate. Higher power dissipation is the main concern with this architecture. A compromise can be found in the folding architecture where the no. of comparators is reduced by using the folding mechanism. In the next section the architectures which can be used for gigahertz range application are described along with the advantages and disadvantages are critically analyzed.
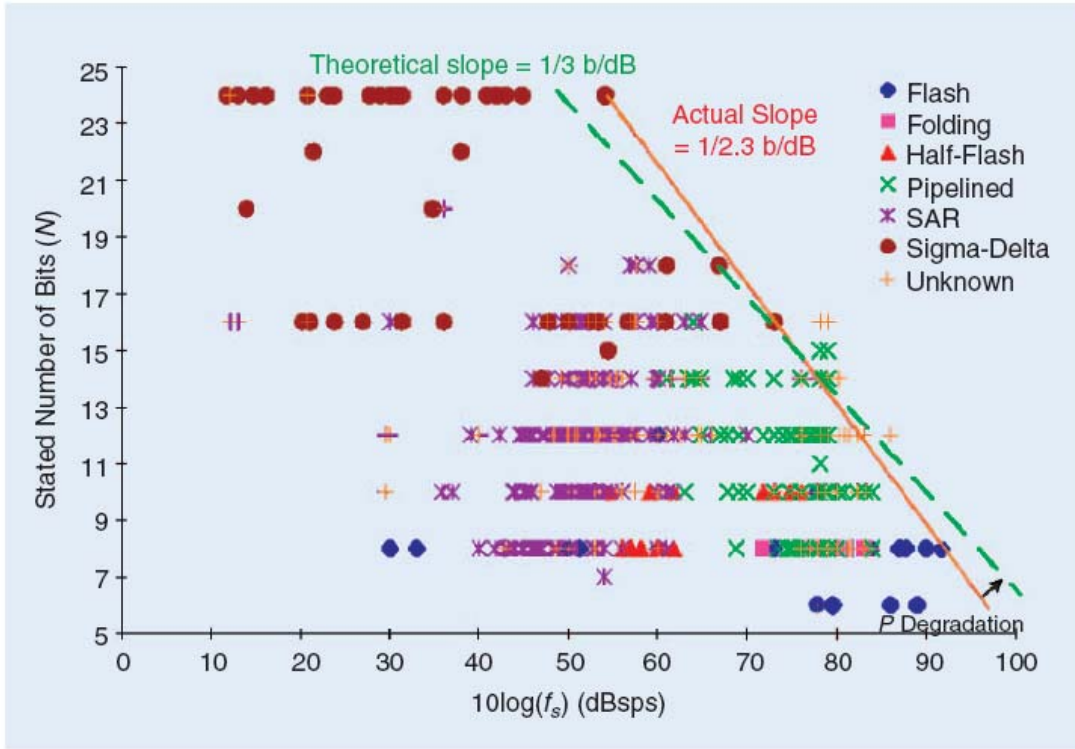
Fig. 2.8. Performance of different ADC architectures [6]

2.5.1. Flash ADC

The flash type ADC is the fastest among the all ADC architectures. A simplified block diagram of this architecture is presented in Fig. 2.9. For a N-bit ADC (2^N-1) number of quantization levels are to be resolved. In this architecture the quantization process is performed by using the same number of comparators (M) as the quantization levels. Thus the maximum amount of the parallelism is employed in this architecture. The reference voltages for the comparators are generated through a resistive ladder. The two ends of this reference ladder are connected with the positive ($+V_{REF}$) and the negative ($-V_{REF}$) references, which determine the full-scale voltage of the ADC. For a given input voltage, comparator outputs from Q_0 to Q_K are logic high and rest of the comparator outputs are logic low. This output pattern is commonly known as thermometer code. The thermometer code is then converted to the binary output by a thermometer to binary encoder.

As the input signal is directly connected to the all comparators, the sampling speed is very fast for this architecture [7],[8]. The speed of the comparators generally limits this sampling rate. The front end sample and hold block can be avoided in this architecture as the sampling operation is directly accomplished by the comparators. Usually the latency is very low for

this kind of architecture typically one to two clock cycles, which makes it useful for feedback applications.
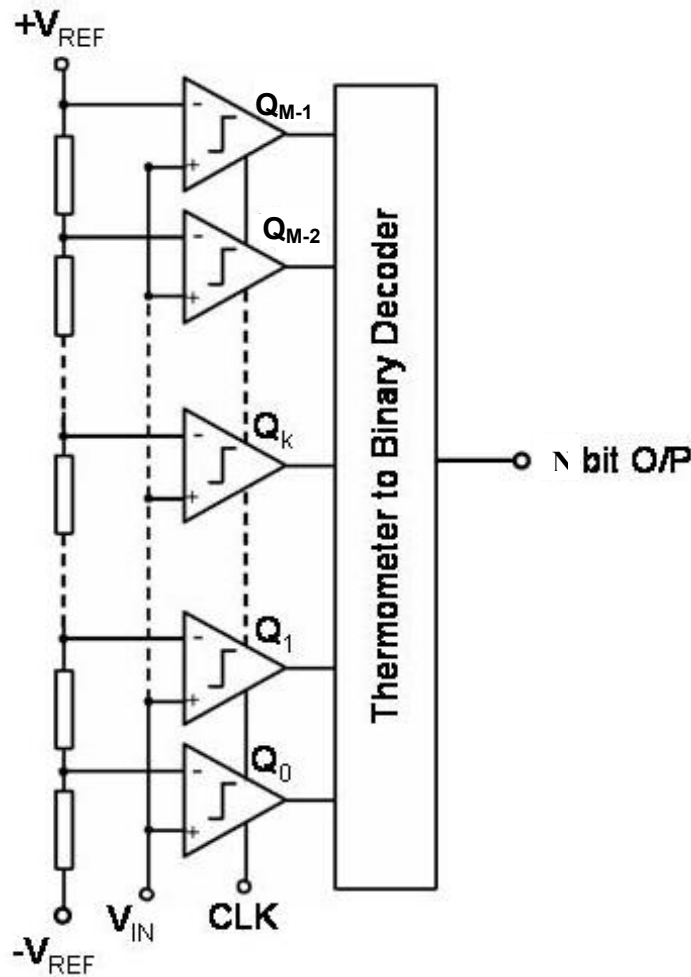
Fig. 2.9. Flash ADC architecture

Certainly the main disadvantage with this architecture is the huge no. of comparators. This number increases exponentially with increasing resolution (N) so does the area and power. Thus in practice this architecture is merely used for resolution more than 8-bits. The large no. of comparators spread over the whole area of a monolithic chip causes higher mismatch among the devices, thus the comparator offsets go high which also restrict the resolution of the full ADC. To overcome this problem large devices can be used but that causes higher capacitive load for the input and reduces the input bandwidth. An alternative solution can be found in the auto-zero comparators [9],[10] where the comparator itself comes with an offset correction mechanism. Generally these comparators have phases. In reset phase the input offset is corrected and in evaluatory phase the real comparison is done.

17

This technique is popular in CMOS technology but cannot be useful for sampling speed in gigahertz range where the bipolar comparators are mainly used.

A well-known technique to improve the static nonlinearity i.e. INL and DNL is to use averaging [11],[12]. In this technique each comparator is preceded by a preamplifier, whose output is coupled to the outputs of the adjacent preamplifiers via a resistive averaging network. As a result, the input signal for a comparator is not produced by its own preamplifier alone, but it is a weighted average of the outputs of the preamplifiers in a small neighborhood. Comparator offset is also reduced by the preamplifier gain and the preamplifier offset is an average of the random offsets of all the amplifiers participating in the averaging process.

Considering all the pros and cons, this architecture is rarely used in high resolution ADCs. The main application is restricted mainly to low resolution disk drive read channel, local area network interface etc. with sampling speed of few hundreds of megahertz. In [13],[14] CMOS flash converters are reported which can work in gigahertz range. The highest sampling speed is reported in [8], where bipolar devices are used for the implementation. Special attention is needed to reduce the clock jitter to enhance the resolution. May be a front end sample and hold can be used to relax the stringent clock jitter requirement in high end applications.

2.5.2. Sub-ranging or Two step ADC

An improvement in the flash architecture can be found sub-ranging or two step ADC. As the name implies the quantization process is performed in two steps. A N-bit ADC is implemented as a combination of two sub ADCs, an M-bit coarse converter which is followed by P-bit fine converter where,

$$N = M + P \tag{2.20}$$

These sub ADCs are implemented with flash architecture. Thus the total no. of comparators is reduced from (2^N-1) to $(2^M + 2^P-2)$. A front-end sample and hold is required to ensure that both of the sub-ADCs are working with the same sampled analog input. After the coarse analog to digital conversion the digital output of the coarse converter is converted back into analog signal by an M-bit sub DAC. The output of the sub-DAC is subtracted from the held analog signal of the sample and held to generate the residual voltage. This residual

voltage is amplified by a factor of 2^M to match the full-scale voltage of the fine ADC with the coarse ADC.
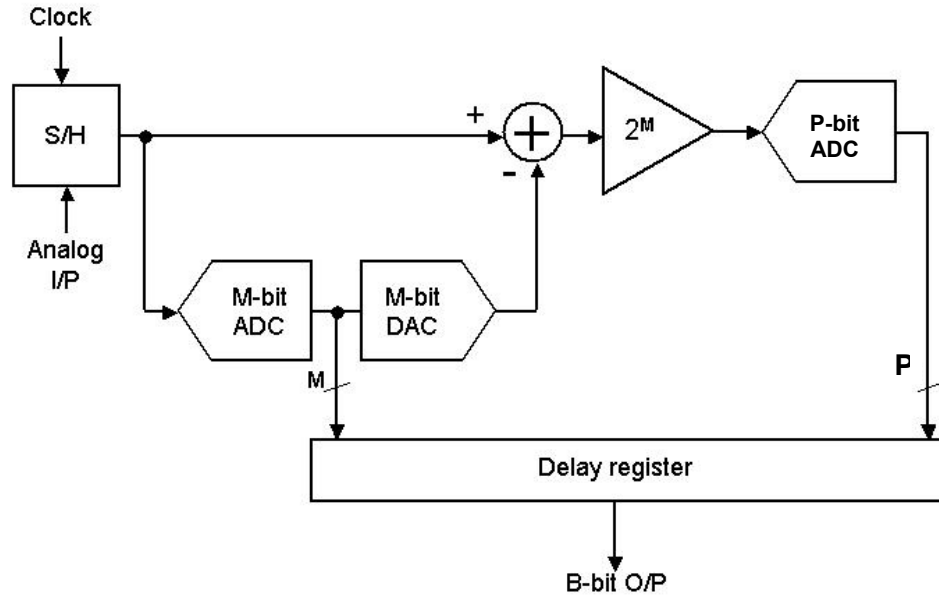
Fig. 2.10. Block diagram of sub-ranging ADC

One of the major drawbacks of this architecture is the non-ideality associated with the comparators. Ideally all of the comparator should have N-bit of accuracy. If the error exceeds the specified tolerance in the coarse converter, an overflow or underflow occurs at the output of the fine converter. To cope with this problem generally redundancy is used in either of the coarse or fine converter. The redundant sign digit (RSD) algorithm [15] similar to the pipeline architecture can be used to relax the comparator accuracy. A vivid analysis of the different error sources for RSD technique is done in [16]. It shows although the comparator accuracy can be relaxed but the accuracy requirement for the sub-DAC remains the same.

In the context of high-speed gigahertz range ADC design this architecture may be useful to reduce the no. of comparators and hence the power dissipation. But the complex residue generation process could appear as the bottleneck. Within a single hold period of the sample and hold three operations (coarse conversion, residue generation and fine conversion) are to be performed which imposes a tough timing constraint. This can be relaxed with inserting another sample and hold in front of the fine converter. But this will introduce another source of error and finally reduce the accuracy of the full converter.

19

2.5.3. Folding ADC

Fig. 2.11. represents a simplified block diagram of folding N-bit ADC. In this architecture the input signal is folded upward or downward after a specified interval as shown in Fig. 2. 12. In this particular example it is ¼ of the full-scale. The output is also same as the input interval. Thus reducing the no. of comparator by the same amount i.e. ¼ compared to the full flash architecture. The folding operation is accomplished by the folder or folding amplifier (Fig. 2. 11). In practice the folding amplifier has a specified amount of gain (in the present example 4) to generate the folding signal with the same full scale rather than the fraction of it. Hence the accuracy requirement of the comparators in the fine ADC is relaxed.
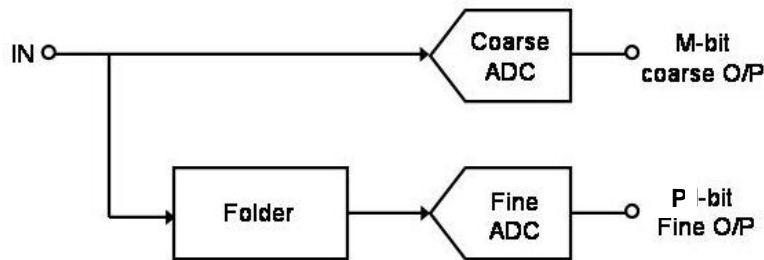
Fig. 2.11. Simplified block diagram of Folding ADC

The concept of folding is similar to the sub-ranging ADC. But here the prior knowledge of sub-range is not required. As a result of it the coarse and fine conversions can be done concurrently and the front-end sample and hold can be avoided which leads to high sampling rate. The output of the coarse ADC is finally used to decode the fine ADC output.
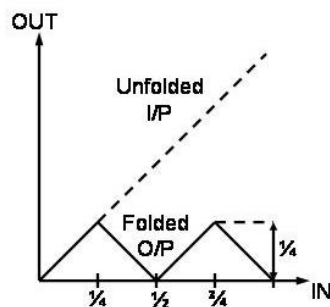
Fig. 2. 12. Principle of folding

In practice, realizing a transfer function of folding amplifier with the triangle wave shape is very difficult, since especially the sharp corners tend to become smoothed due to the limited bandwidth. This problem can be solved by producing several versions of the folded signal; each shifted a different amount of input voltage in the x-direction, and using only the

linear part of each curve. This is illustrated in Fig. 2. 13. where five nonlinear curves are used instead of one linear one. The linear portion around the zero crossings of each curve is utilized for comparison. All the comparators responsible for detecting the signal in this range are connected to the circuit producing that particular curve. Often, the number of curves is increased up to the point where they equal the number of comparators. As a result, there is only one comparator per curve and it only has to detect the signal zero crossing, making the linearity of the curve unimportant.
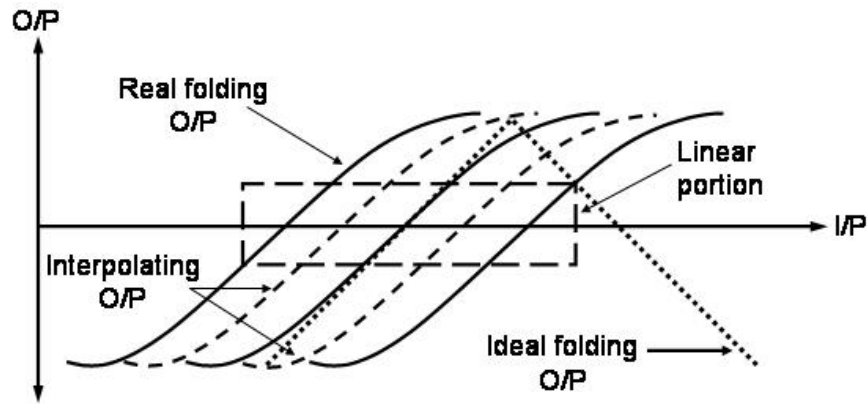
Fig. 2. 13. Folding signal generation

The folding amplifier is a complex block, but a large no. of folding block is not used. Instead interpolation technique is used. The interpolation curves are achieved by shifting the real folding signals in the y-direction (shown by the dashed lines in Fig. 2. 13). The interpolating signals do not represent the actual folding signal rather it represents the valid information only in the vicinity of the zero crossing. The main advantage of this technique is the fact that the process can be implemented by simple resistive ladders. Thus ideally a large no. of interpolating signals can be used without increasing the complexity. A block diagram of the folding ADC used in practical applications is shown in Fig. 2. 14. In this example four folding amplifiers are used. The coarse ADC determines the sub-range for the folding amplifiers. M-bits are resolved from the fine ADC. The folding amplifiers accomplished with the resistive interpolators define the zero crossings for the each comparator in the fine ADC. The outputs of the coarse and fine ADCs are fed to a decoder to generate N-bit output.

A possible implementation of the folding amplifier with bipolar devices is presented in [17]. The folding amplifiers are implemented with open loop parallel emitter couple pairs. As a result it becomes suitable for the high-speed applications. High g_m of the bipolar transistors guarantees the higher linearity of the folding amplifiers. The main disadvantage of the folding amplifier is the fact that the output frequency of the block is the product of the input

21

frequency and the no. of folding used (see Fig. 2. 12). Sometimes this comes as the decisive factor for the input bandwidth. This problem can be bypassed by using a front end sample and hold. The resolution of this architecture is limited to 8-10 bits. In [17] an 8-bit resolution is achieved with 2GHz of sampling rate.
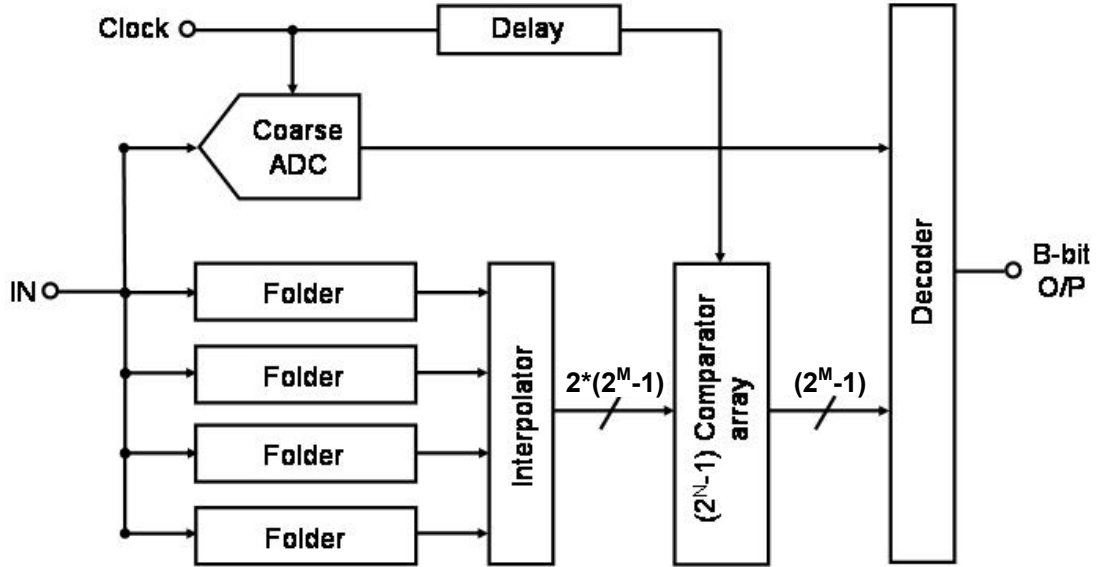
Fig. 2. 14. Folding interpolating ADC Architecture

2.5.4. Time Interleaved ADC

A simplified block diagram of the time interleaved ADC is presented in Fig. 2. 15. In this architecture, M no. of ADCs are used in parallel (known as channel) to enhance the sampling rate to M times of an individual ADC. Each ADC works on every M^{th} sample value. At the output a multiplexer is used to select the output of the proper ADC to generate a single bit stream at the full sampling rate.

Up to a certain resolution the component mismatches are within the tolerance. But with increasing resolution severe problem with static characteristics occurs in this architecture due to the gain and offset mismatch of the different channels assuming that each channel works with the same linearity. The offset error can be overcome easily by using mixed mode [18] or full digital calibration [19] technique. Calibration of gain requires more complex circuitry [20].

Certainly the main problem in the dynamic behavior occurs due to the clock skew for different ADC channels. This can happen due to the clock generation circuitry or the possible propagation delay mismatch among the different sampling circuitry. One of the favorable solutions would be to use a front-end sample and hold working with the full

sampling rate. But in high frequency sampling regime, with the increasing no. of parallel channels it becomes very difficult to drive the large capacitive load. The clock skew problem can also be solved by digital post-processing. But it requires an accurate measurement of the clock skew in the sub-pico second range.
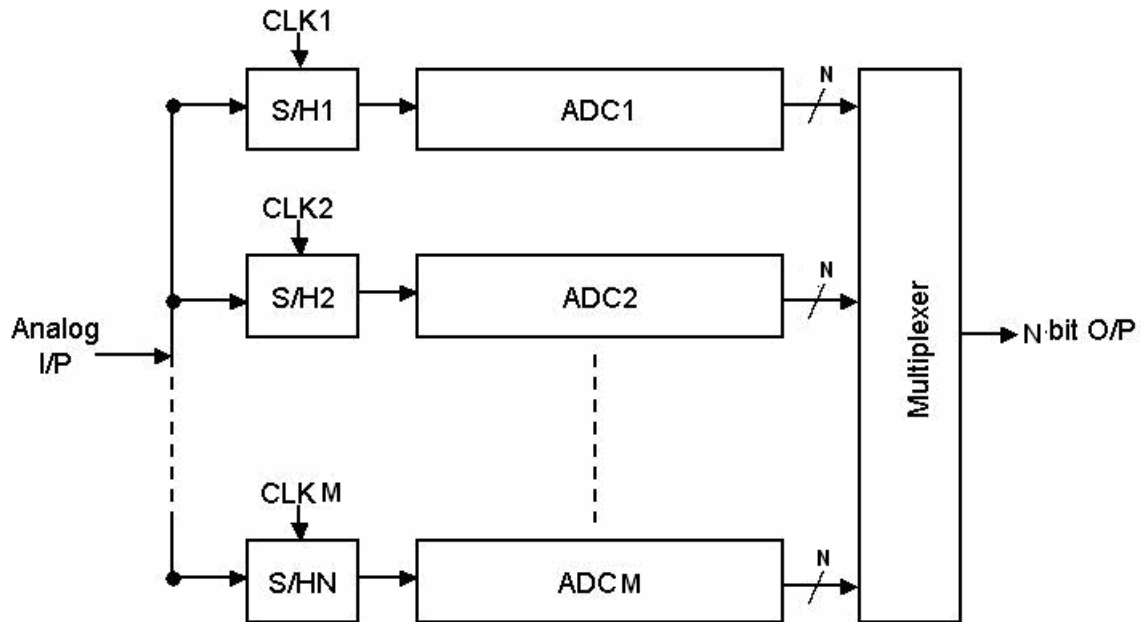
Fig. 2. 15. Block diagram of time interleaved ADC

Until now the best performance using this architecture is reported in [21] which in fact represents the best performance in terms of sampling speed and resolution in SiGe technology. In that work eighty parallel current mode pipeline ADCs were used. A complex DLL based clock generation scheme is employed to achieve lower clock skew among the blocks. An on-chip 1MB memory is used to store the output of the parallel channels and the final digital output is achieved after performing digital post processing on this data. Thus for real time application it is not suitable.

2.6. Conclusions

In this chapter the basic quantization process is described. The mid-riser quantizers are found to be most commonly used quantization method. The static and dynamic errors associated with the quantization process are defined. The physical error sources which define the limit for the ADCs in terms of resolution and sampling rate are identified as the input

referred thermal noise, the aperture uncertainty in the sampling process and the comparator ambiguity.

Different ADC architectures which can be used for gigahertz range sampling are discussed. The flash type architecture is found to be the fastest but power dissipation is highest in this architecture. An alternative can be found in the time interleaved ADCs. In fact the best performance is achieved by using this architecture. But it comes with a large amount of digital post processing overhead which makes it unattractive for real time applications. The compromise in resolution, speed and power can be found in folding architecture where the coarse and fine conversion can be done concurrently. But the bandwidth limitation of the folding amplifier may come as a bottleneck. This can be overcome by front end sample and hold.

Chapter 3
Design of Multi-GHz ADC Components

_____

3. 1. Introduction

In this chapter the design of two main ADC components are presented. These components can be used as standalone systems and as well as the building blocks for a complete ADC system. In the first part of this chapter the design of the track and hold amplifier (THA) is presented. For any high-speed high resolution ADC the front end THA comes as the most critical component. The error introduced in this block cannot be suppressed by post processing of the sampled signal.

The main operation of THA can be divided into two phases. In the first phase THA follows the input signal. Then it goes to the next phase in which it holds the sampled value for a finite time. In general these two phases have the same time intervals.

For high speed application open loop architectures are commonly used. Unlike the closed loop THAs, the global feedback from the input to the output is not present in this kind of topology. As a result of it the linearity is not very high. The nonlinearity increases with the increase in the input voltage range. On the other hand the quantization process can be efficiently done if the input range of the THA is high. In some of the applications almost 2Vpp differential input is required for the quantization process [17]. In modern state-of-the-art SiGe technologies with collector-emitter breakdown voltage (BV_{CEO}) around 2V, such a high swing is difficult to obtain due to the nonlinearity inherent to all stages operating close to their swing limit. In this chapter two different open loop THA architectures are proposed which are capable to work with high input swing at the sampling rate of 10GS/s.

In the second part of this chapter the design of a 20GS/s comparator is presented. The comparator is implemented with the open loop architecture. In measurement it shows 5.8-bit of accuracy with 70mW of power dissipation.

The chapter is organized as follows, in section 3. 2 the definitions different parameters of THA performance matrices have been presented. Brief reviews of the most commonly used

THA architectures are presented in section 3.3.The design technique of two different THAs for high input swing is described in section 3. 4. In section 3. 5 a new architecture double sampling THA architecture is presented. The experimental results of the implemented THAs are presented in section 3. 6. In section 3.7 design of an open loop comparator is presented which is followed by the measurement results in section 3.8. Finally conclusions are drawn in section 3.9.

3. 2. Performance Matrices for Track and Hold Amplifier

The basic track and hold operation is divided into two phases. In the first phase THA works as a unity gain amplifier and follows the input signal. In the second phase the THA holds the track voltage. A simplified functional block diagram of a THA is shown in Fig. 3.1. A unity gain input buffer is used to isolate the sampling circuit from the outer-world. The main track and hold function is accomplished by the sampling switch. This switch is controlled by a clock signal and the input analog signal is stored across the hold capacitor C_H. Finally an output buffer is used to isolate the C_H from the external load.
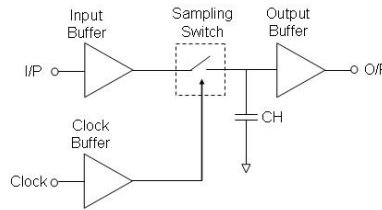
Fig. 3.1. Functional block diagram of THA

The performance of THA can be characterized by a number of parameters. The terminologies and definitions used to characterize THA vary with different manufactures. In this section mostly acceptable performance parameters are defined. As mentioned earlier in the track mode the THA works as a unity gain amplifier. Thus in this mode the THA is characterized by the same parameters like analog amplifier e.g. offset, gain, slew rate, bandwidth, nonlinearity, harmonic distortion and settling time. In Fig. 3.2 the terminologies related to THA timings in both the track and hold phases are depicted.

The acquisition time is the time interval, during which the THA must remain in the track mode to enable the circuit to accurately replicate the input signal, thereby ensuring that the subsequent hold mode output will lie within a specified error band of the input level that existed at the track-to-hold transition (after gain and offset effects have been removed). The

remaining time duration of the track mode exclusive of acquisition time is called the track time during which the THA output is a replica of its input.

The settling time is defined as the time duration between the beginning of track to hold mode transition and the time when THA output is settled down within a specified error band of the final hold value. The remaining time in the hold mode can be used for post processing e.g. analog signal processing, analog to digital conversion.
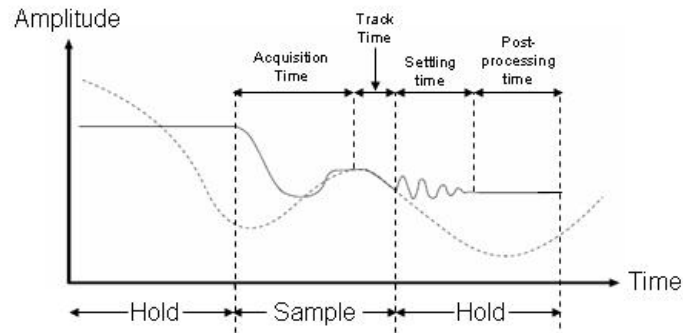
Fig. 3. 2. Track and Hold terminologies

The track to hold transition determines many aspects of T/H performance. The delay time is the time elapsed from the execution of the external hold command until the internal track-to-hold transition actually begins. In practical circuits this switching occurs over a non-zero interval called the aperture time measured between initiation and completion of the track-to-hold transition. Practical circuits do not exhibit precisely the same time period for each sample. This random variation from sample to sample is caused by phase noise of the incoming clock signal and further exacerbated by electronic noise within the T/H itself. The standard deviation of the sample period is termed the aperture jitter and limits amplitude resolution in A/D conversion.

In Fig. 3.3 different error sources in the hold mode are being shown. During the transition from the track to hold mode an error in the hold voltage is introduced which is known as pedestal error. This error stems from the charge injection of the sampling switches.
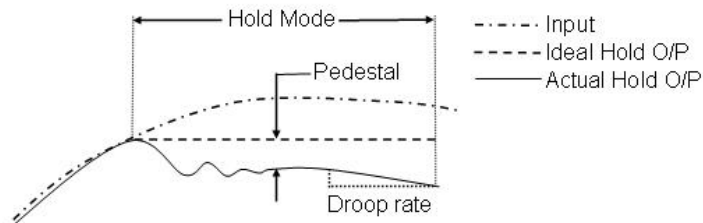
Fig. 3. 3. Hold mode characteristics

Due to the leakage current from the hold capacitor the hold mode output decays with a constant rate. This rate of decaying is known as droop rate. This error can be reduced by

27

differential designs. The parasitic coupling from the input to the output in the hold mode is defined as hold mode feedthrough.

3. 3. Open Loop THA Architecture Review
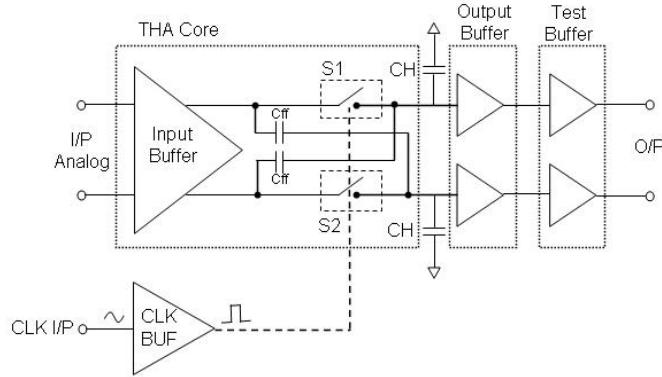
Fig. 3.4. Block diagram of open loop THA

In a high-sampling rate regime, the open loop architecture is suitable choice [[22], [23]. Closed loop THA are much slower considering the fact that the feedback loop has higher time constant and settling time. In the Fig. 3.4 the block diagram of open loop THA is presented. It has three main sub-blocks. A unity gain amplifier is used as the input buffer. This followed by a pair of sampling switches. The main sampling operation is accomplished with these switches. For open loop applications diode bridge switches [25] can be used. But this comes with higher voltage headroom requirements. An improvement is proposed in [26] but it requires a complex pulse shaping circuitry to control the switch. The most commonly used sampling switch is known as the switch emitter follower (SEF). In most of the well-known open loop architectures different SEF topologies are used. Although the core structure remains the same. Two cross-coupled capacitors (C_{ff}) are used to reduce the hold mode feedthrough.

After the sampling switches another unity gain amplifier is used to isolate the hold capacitor (C_H) from the external load. Sometime an additional buffer (test buffer in Fig. 3.4) is also included to drive the external 50Ω load. For high-speed application an external sinusoidal signal is used as clock and an on-chip limiting amplifier is used to generate the clock signal.

3. 3. 1. Open Loop THA with Switched Emitter Follower

The most commonly used open loop THA architecture is presented in Fig. 3.5 [22]. A differential pair with emitter degeneration resistors (R3, R4) is used as an input buffer. For better linearity, resistors in series with diode-connected transistor loads are used. The non-

28

linear voltage-to-current conversion by the input transistors due to the base-emitter voltage modulation is compensated by the current-to-voltage conversion through the load resistance (R1, R2) and the diode connected load (Q3, Q4). This configuration results in good linearity in low frequency range. With the increasing input frequency the possible delay mismatch among the input (Q1, Q2) and the load transistors (Q3, Q4) affect the compensation process. Beside that the input swing in this configuration is restricted due to the fact that the large input may cause breakdown in Q1 and Q2, which results in nonlinear current to voltage conversion.
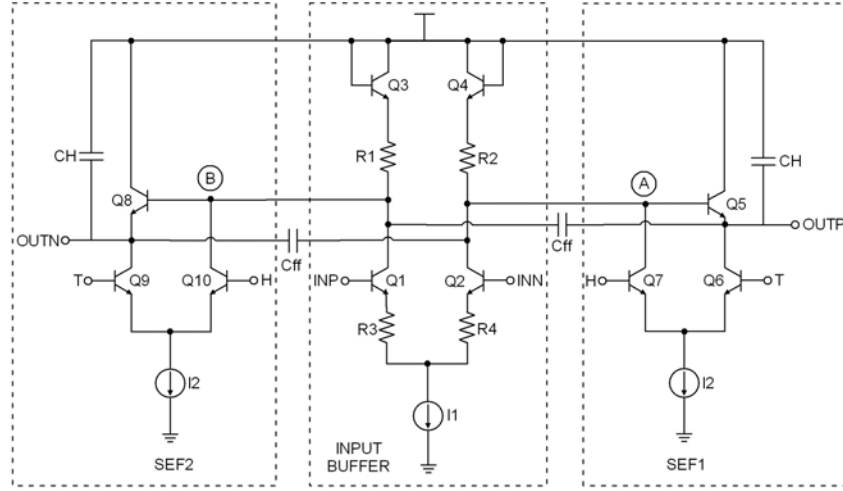
Fig. 3. 5. Open loop THA implementation

The sampling switch is generally known as switch emitter follower (SEF). It consists of three transistors and a tail current source. The emitter coupled transistors (Q6, Q7 and Q9, Q10) act as current switch and C_H is the hold capacitance. When the track signal (T) goes high Q5, Q8 work as emitter followers and Q6, Q9 appear as cascode transistors. In the hold mode the tail current is switched through Q7 and Q10. As a result the voltage at nodes A and B become lower which turn off Q5 and Q8 to store the samples voltage across the hold capacitor. During the hold mode the tail current (I2) flows through the resistors R1 and R2. It may pull down the voltage of node A and B to such an extent, which drives Q1 and Q2 into saturation. This drastically reduces the speed of operation. So R1, R2 and the tail current (I2) have to be optimized.

In the hold mode some portion of the input signal is coupled to the hold capacitance C_H through the parasitic base-emitter capacitance of Q5 and Q8. This leakage is known as hold mode feedthrough. In differential design ideally this feedthrough is equal and opposite for the two differential paths. To compensate this error the feedthrough voltage is cross-coupled to

29

the hold capacitances (C_H) by C_{ff} capacitors. The C_{ff} is actually series and parallel combination of four diode-connected transistors.

3. 3. 2. Improved Open Loop THA Architecture
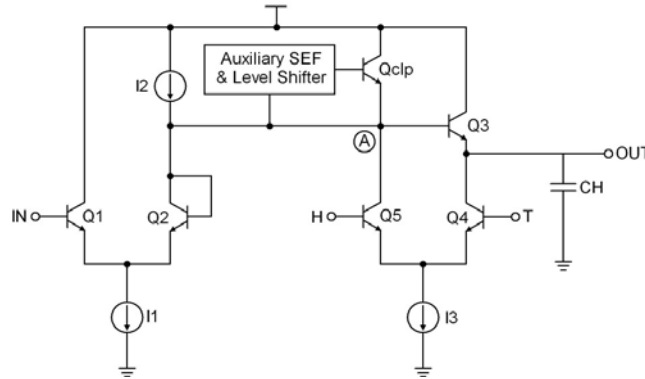
Fig. 3. 6. Improved Open Loop THA

An improved open loop THA [23] is presented in Fig. 3. 6. In this figure only one of the pseudo differential paths is shown. A differential pair in unity feedback configuration is used as an input buffer and current source (I2) is used as a load. At the output node of the input buffer (A) the DC level is similar to the input DC level. Commonly a pMOS current source is used to implement I2. The higher open loop gain of the differential pair results in good output linearity. The output resistance of the differential pair will be near to $1/g_m$ as long as the pMOS current source has higher output impedance. But in reality with the increase of input frequency the output impedance of the pMOS current source decreases and the parasitic capacitance associated with the pMOS transistor limits the sampling rate of the THA. An improvement is proposed in [27] to overcome the problem associated with the finite impedance of the pMOS current source for high frequency of operation. In this solution an inductive degeneration is used at the source node of the pMOS to achieve comparatively higher impedance for high frequency operations.

The core architecture of the sampling switch is similar to the SEF discussed in section 3. 3.1. The performance of the sampling switch is enhanced by employing a clamping transistor (Q_{clp}) at the input node (A) of the switch. In the track mode transistor Q3 works as an emitter follower and tracks the input signal. The current switching for the track to hold transition is done in the same fashion like the conventional SEF, i.e. by the means of the emitter-coupled pair Q4 and Q5.

The hold mode operation substantially differs from the conventional SEF. In this mode the I3 current flows through Q5 and Q_{clp} appears as an emitter follower. If I3>I2 then Q_{clp} pulls down the potential of node A by the amount of base-emitter voltage (V_{be}) of it. The DC level of the base of Q_{clp} is kept at the same potential that of the input (V_{DC}) by the means of a level shifter. Assuming the base-emitter voltage of Q1 and Q_{clp} are same, in this mode the DC node voltage of node 'A' falls down by an amount of V_{be} compare to the track mode. Thus the base-emitter diode of Q3 is turned off and C_H is isolated from the input buffer. Additionally the base-emitter voltage of Q2 becomes zero, which restricts it to follow the input signal. The load current I2 is sink through Q5.

The transient waveform of node A is plotted in the Fig. 3. 7. It is assumed that the base of Q_{clp} is directly connected to the input instead of an auxiliary SEF. In the hold mode level shifted input appears at node A and coupled to the C_H through the parasitic base-emitter capacitance of Q3. This feedthrough can be reduced by feeding back the level shifted version of the hold voltage to the base of Q_{clp}. This will produce a flat top signal at the node A. The feedback signal is not taken directly from the C_H instead it is produced by an auxiliary sampling switch, which has the same architecture like the conventional SEF.
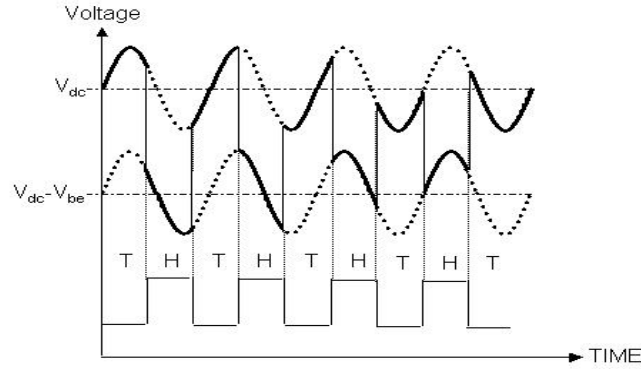
Fig. 3. 7. Transient waveform at the input node (A) of the sampling switch

3. 4. Implementation of Open Loop THA

In this section the implementation of two open loop THAs are presented. One is implemented with all npn transistors whereas the other uses the complementary npn and pnp transistors. The core-sampling switch is implemented with the simple SEF described in the section 3.3.1. The main difference in these implemented THAs can be found in the input buffer. In the npn THA, a simple cascode amplifier is used as the input buffer whereas an

improved npn pnp emitter follower is used in the other THA. The detail design considerations and the main error sources in different sub-blocks are discussed in the following sections.

3. 4. 1. Implementation of Input Buffer

In section 3.3.1 nonlinearity associated with the emitter degenerated differential input buffer is discussed. The nonlinearity increases with the input voltage range due to the nonlinear output characteristics of the transistors near to the collector-emitter breakdown region. For high-speed applications the transistors need to be biased in the high current density region to achieve higher f_T and for the better switching speed the input transistors are to be biased with higher collector-emitter voltage (V_{CE}). This imposes a limitation in the input range, particularly in the sub-micron technologies as the collector-emitter breakdown voltage (BV_{CEO}) is diminishing with the feature size of the transistors.   In the present work THAs are implemented in a commercially available 0.25μm BiCMOS technology. The npn HBTs are having f_T / f_{max} of 190GHz/190GHz and BV_{CEO}=2.0V.

In this section two different variants of input buffers are presented which can provide acceptable linearity with higher input range. The nonlinearity problem associated with low BV_{CEO} is overcome by the use of a complementary emitter follower and cascode input stage.

3. 4. 1. 1. Complementary Emitter Follower

The input buffer used in [22] can be very easily replaced with a pnp emitter follower. This could be advantageous in enhancing the bandwidth of the THA. Due to the fact that, for a given bias current and the load the bandwidth of the emitter follower is inherently higher than the differential pair. Secondly it provides a well-defined gain near to unity, which is less dependent on the process parameter variations. In Fig. 3.8a a simple pnp emitter follower is presented. The output voltage is given by,

$$V_{out} = V_{in} - V_{BEP} \qquad\qquad (3.1)$$

Where V_{in} and V_{out} are input and output voltages respectively and V_{BEP} is the base-emitter voltage of the pnp emitter follower.

For a large input swing two main sources of nonlinearity can be identified, which reduce the linearity of the emitter follower. The input signal dependent variation of V_{BEP} from its quiescent value (V_{BEPQ}) comes as a distortion at the output. If it is assumed that the emitter follower transistor is biased with the collector current, I_{CQ} then the incremental output voltage (ΔV_{OUT}) can expressed be as,

$$\Delta V_{OUT} = \Delta V_{IN} - \Delta V_{BE} \tag{3.2}$$

where, ΔV_{IN} is the incremental input voltage. For simplicity the base emitter voltage V_{BEP} is replaced by V_{BE} and ΔV_{BE} is incremental error due to the change in input. The equation 3.2 can be expressed in terms of quiescent collector current (I_{CQ}) and the incremental error (ΔI_C),

$$\Delta V_{OUT} = \Delta V_{IN} - V_T \ln\left(1 + \frac{\Delta I_C}{I_{CQ}}\right) \tag{3.3}$$

According to equation 3.3 biasing the emitter follower with high collector current can reduce the fractional error the V_{BE} modulation.

(a) (b)
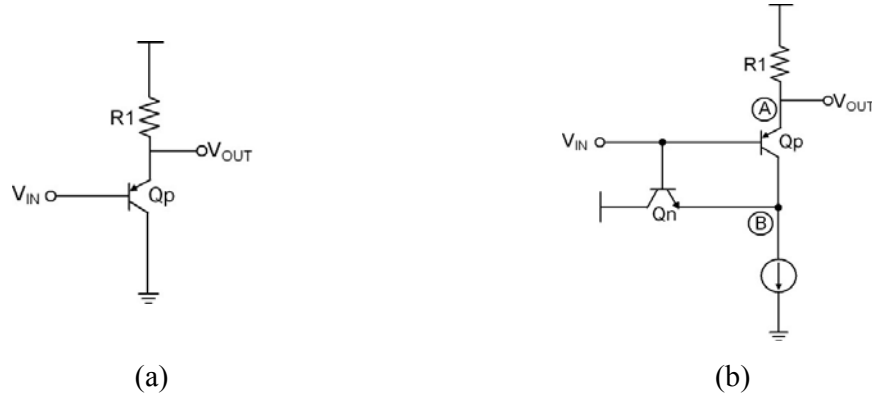
Fig. 3. 8. (a) Simple pnp emitter follower (b) npn-pnp emitter follower

The second source of nonlinearity associated with the simple pnp emitter follower is the nonlinear output characteristics near BV_{CEO}. The emitter follower that has to be used as the input buffer is generally bias at a higher collector-emitter voltage (V_{CE}) to ensure that in the hold mode when the output node voltage is pulled down transistor Q_p is not pulled into saturation. Therefore the output swing of the input buffer is reduced.
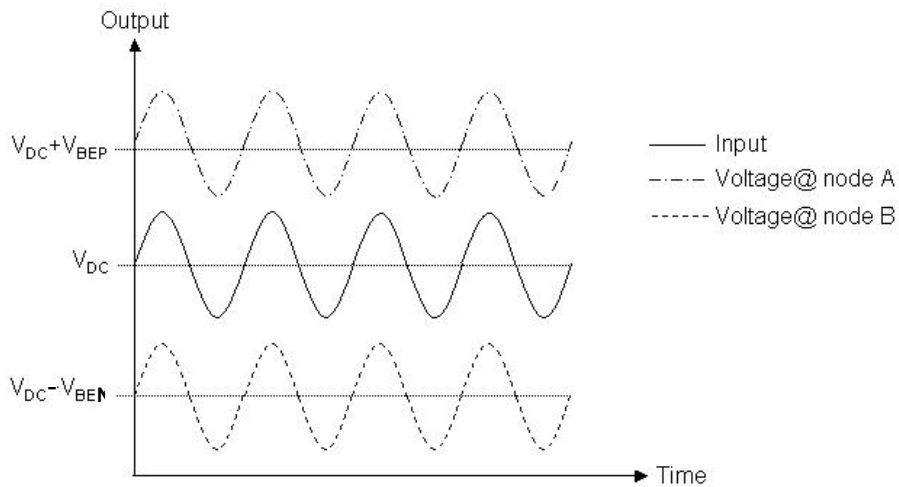
Fig. 3. 9. The voltage wave forms at different nodes of npn-pnp emitter follower

This problem of increasing V_{CE} with the input amplitude can be solved by the improved emitter follower structure proposed in Fig. 3. 8b [28]. In this proposed emitter follower structure the pnp transistor (Qp) is used as the main device and a feedforward path to the collector node of Qp (B) is provided by an auxiliary npn emitter follower (Qn).

In Fig. 3. 9 the collector and emitter node voltages of Qp are shown for a given input DC level (V_{DC}). The Qn reproduces a replica of the input signal with a level shift of $-V_{BEN}$ at node B (See Fig. 3.8). On the other hand the at the collector node of Qp (A) the input is shifted by $+V_{BEP}$. If the difference of delay between Qp and Qn are assumed to be small then the collector-emitter voltage of Qp is fixed to $V_{CEP}=V_{BEP}+V_{BEN}$ which results in a better linearity.

In Fig. 3. 10 the simulated third order harmonic power of the pseudo-differential input buffer with pnp-npn EF is plotted for 1GHz of input signal with different amplitude, in comparison with the conventional npn differential stage. Both the buffers are optimized for the 2Vpp input differential signal with the same power supply. At about 1.0V of differential input voltage, the distortion of the standard input stage starts to grow rapidly, while the input buffer with pnp emitter follower keeps below –50dBc up to 2V. At 2V, the distortion of the proposed circuit is about 16.8dB less than that of the conventional input stage, corresponding to an increase in ENOB by 2.5 bits.
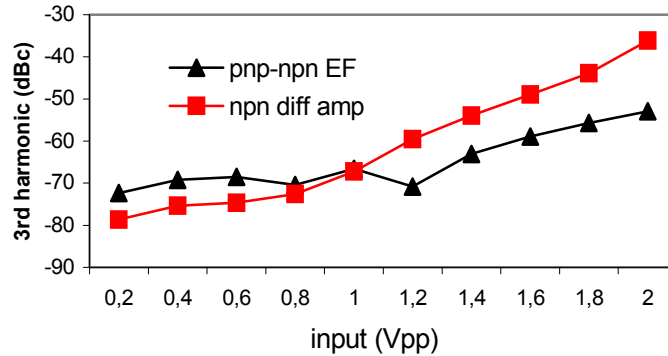
Fig.3. 10. 3rd harmonic power of npn pnp emitter follower input buffer

3. 4. 1. 2. Cascode Input Buffer

The input buffer mentioned in the earlier section uses a pnp transistor as the main device. In commercially available BiCMOS processes the pnp device can rarely be found. In this section an alternative to the npn pnp emitter follower is proposed. A cascode amplifier is a

well-known structure for better linearity and higher bandwidth. But it has not yet been used as an input buffer of the open loop THA.
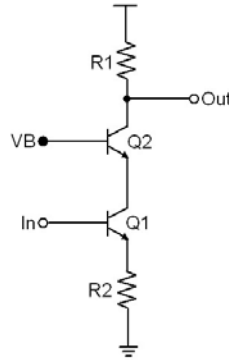
Fig. 3. 11. Cascode input buffer

The cascode amplifier, which is used as an input buffer [29], is shown in Fig. 3. 11. The gain of this amplifier is determined by the ratio of the resistances R1 and R2. The common base transistor Q2 provides almost a fixed voltage at the collector node of the main common emitter transistor Q1. The base voltage of Q2 (V_B) can be optimized to maximize the input range of Q1 without pulling it into weak collector-emitter breakdown region. On the other hand the output swing can be high enough. Owing to the fact that collector-base breakdown voltage (BV_{CBO}) of the transistor is much higher than BV_{CEO}.

A linearity comparison similar to the npn pnp emitter follower has been performed for 1GHz input sinusoidal. In Fig. 3. 12 the 3rd harmonic power for different input amplitudes plotted for the pseudo-differential cascode input buffer and the conventional emitter degenerated differential input stage. In the conventional input buffer the 3rd order harmonic power increases with input level at an average rate of 23.8 dB/V whereas in the cascode amplifier increases at 17dB/V. For 2Vpp differential input the difference is approximately – 20dBc, which corresponds to an improvement of 3.02 effective number of bits (ENOB).
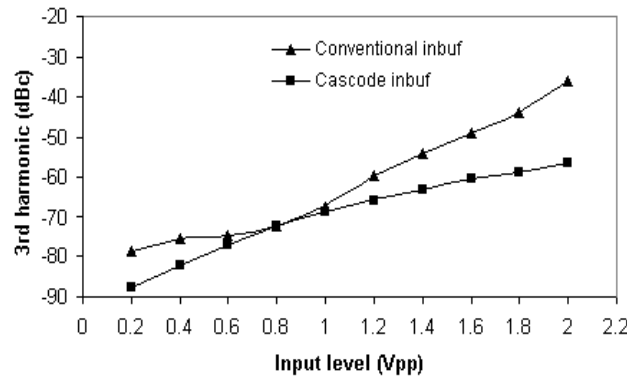
Fig.3. 12. 3rd harmonic power of cascade input buffer

35

3. 4. 2. Implementation of Switched Emitter Follower

The switched emitter follower (SEF) is the most important part of the open loop THA. In section 3. 3 two different kind of SEF architectures are presented. In this present work the simple SEF (Fig. 3. 5) is preferred over the improved SEF structure, as the complex switching requires an additional pair of auxiliary sampling switches. In high-speed application it comes with additional power consumption overhead.

A simplified schematic diagram of the SEF is presented in Fig. 3. 13. As mentioned earlier, in the track mode the SEF works as an emitter follower and the performance of the switch in this mode can be expressed in terms of gain, offset and THD. The transition form the track to hold mode comes with a number of non-idealities. Those errors can be categorized broadly into two categories; the timing related and the amplitude related errors. The timing related errors stem from the finite aperture time and the aperture jitter whereas the amplitude related errors are characterized by the pedestal error and the coupling of input signal to the hold capacitance in the hold mode i.e. the hold mode feed through. In this section, these error sources are described and their impacts in the designing of SEF are discussed.
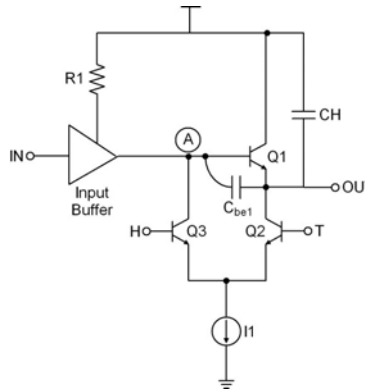
Fig. 3. 13. Switch emitter follower

3. 4. 2. 1. Aperture Time

In Fig. 3. 13. one of the differential (or pseudo-differential) paths of the THA is shown. In the track mode transistor, Q1 works an emitter follower and charges the hold capacitance C_H. The equivalent approximation at this mode is shown in Fig. 3.14.

The diode D1 is used to model the base-emitter voltage of the transistor Q1. During the transition from the track to hold mode the bias current I1 is switched from the transistor Q2 to

36

Q3 pulling the node A to a lower voltage and turn off the base-emitter diode of Q1. This transition takes a finite time τ_a which is known as the aperture time. The value of τ_a depends on the time constant at node A and the bias current I1. Under the assumption, the collector current switches from the transistor Q2 to Q3 linearly during this time, then the equivalent resistance of the diode D1 increases exponentially to infinity. Thus, the error charge accumulated across the hold capacitance can be expressed as,

$$Q_a = \int_{t_0}^{t_0+\tau_a} \frac{V_A(t) - V_{out}(t)}{Z_1(t)} dt \qquad (3.4)$$

Where, V_A is the node voltage at A which is identical with the input voltage. $Z_1(t)$ is the instantaneous impedance of the base emitter diode of the transistor Q1. As the voltage across the base-emitter terminals are changing so the base-emitter impedance is also a function of time.
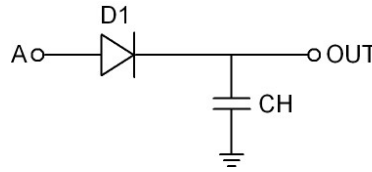
Fig. 3. 14. SEF approximation in the track mode

Even though transistors Q2 and Q3 are arranged in a differential stage to make the switching process more symmetric, they operate at different collector voltages causing some timing difference $\Delta \tau_a$. This produces a charge offset $I_1 \Delta \tau_a$ during the transition [24]. Another source of error is the clock signal (T/H) coupling into C_H through the base-collector capacitance (C_{bc2}) of Q2. Finally, the total error charge introduced due to the finite aperture time (τ_a) is given by,

$$Q_{aperture} = \int_{t_0}^{t_0+\tau_a} \frac{V_A(t) - V_{out}(t)}{Z_1(t)} dt + \Delta \tau_a \cdot I_1 + C_{bc2} \cdot V_{CLK} \qquad (3.5)$$

where the V_{clk} is the amplitude of the sampling clock and C_{bc2} is approximated as constant capacitance.

3. 4. 2. 2. Pedestal Error

While the SEF is conducting during track mode, the emitter follower transistor Q1 stores charge in the base emitter capacitance (C_{be1}). After the SEF switches to hold mode and all transients being settled down, Q1 conducts no current. The difference in charge stored during

track mode and hold mode is therefore expelled from C_{be1} during the turn-off transient. A fraction (η) of charge (Q_{inj}) is injected onto hold capacitor imparts an output voltage perturbation called hold step or hold pedestal which can be expressed as,

$$Q_{inj} = \eta \cdot \int\limits_{V_A - V_{out}}^{(V_A - I_1 R_1) - V_{out}} C_{be1} \, dV \qquad (3.6)$$

where, V_A is the dc level at node A in the track mode and in the hold mode it becomes ($V_A - I_1 R_1$). The base-emitter capacitance C_{be1} varies with time during the SEF transition from track to hold mode. In this analysis C_{be1} is assumed to be the equivalent base-emitter capacitance of Q1 during track to hold mode transition.

In the differential design, the imbalance in between the injected charge of the differential paths determines the effective charge injection (Q_{eff}), which the corresponding error voltage is ΔVp. Assuming that a perfect matching in between the differential path, ΔVp is given by,

$$\Delta V_p = \frac{Q_{eff}}{C_H} = \frac{1}{C_H} \cdot \eta \cdot \int\limits_{-V_{out}}^{V_{out}} C_{be1} dV \qquad (3.7)$$

where V_{out} and $-V_{out}$ are the differential out voltage in the hold mode and C_H is the hold capacitance. η depends on the instantaneous impedance of the base and emitter node of the transistor Q1. The effective pedestal can be reduced by the bigger hold capacitance.

3. 4. 2. 3. Hold Mode Feedthough

When the SEF is in the hold mode, the emitter follower transistor Q1 presents finite impedance. Due to the parasitic leakage of the input voltage through the base emitter capacitance C_{be1} (see Fig. 3. 13) the hold voltage across the capacitance is perturbed, this effect is known as the hold mode feedthrough. The hold mode feedthrough (A_f) is given by the following equation [22],

$$A_f = \frac{C_{be1}}{C_{be1} + C_H} \qquad (3.8)$$
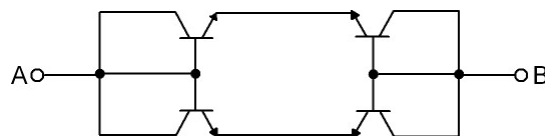
Fig. 3. 15. Hold mode feedthrough compensation capacitor

$$A_c = \frac{C_{be1}}{C_{be1} + C_H}\left(1 - \frac{C_{ff}}{C_{be1}}\right) \tag{3.9}$$

This hold-mode feedthrough can be reduced by adding two feedforward capacitors (C_{ff}) as shown on Fig. 3. 5. The charge dump of these capacitors is of opposite sign to the charge dump of the base-emitter capacitances of the switching transistors. The compensated hold-mode feedthrough (A_c) is now given by,

Complete cancellation of the hold-mode feedthrough would require C_{ff} being identical to the base-emitter capacitance. The feedthrough capacitor (C_{ff}) is realized using a series-parallel construction of four diodes (Fig. 3. 15) [22]. In reality, the device mismatch of the HBTs limits the cancellation.

3. 4. 2. 4. Aperture Jitter

Random variation in the sampling period due to different electronic noise and the phase noise of the input clock is known as the aperture jitter. This error translates into the effective amplitude error of the THA and reduces the resolution. In practical systems with the presence of the clock jitter, the sampling period (T) can be expressed as,

$$T = T_{nom} + \tau_{jitter} \tag{3.10}$$

where, T_{nom} is the ideal time period and τ_{jitter} is the random error due to the clock jitter. The clock jitter is generally expressed as a random variable, which has a zero mean normal probability distribution function with the variance of σ_τ such that,

$$P(\tau_{jitter}) = \frac{1}{\sqrt{2\pi\sigma_\tau^2}}\exp\left(-\frac{\tau_{jitter}^2}{2\sigma_\tau^2}\right) \tag{3.11}$$

Assuming a input sinusoidal V_{in}= $A\sin(2\pi f_{in}\cdot t)$= $A\sin(\omega_{in}\cdot t)$, The average noise power contribution (σ_{jitter}^2) due to the clock jitter can be estimated as,

$$\begin{aligned}\sigma_{jitter}^2 &= \varepsilon\left\{\frac{1}{T_{in}}\int_0^{T_{in}}(\Delta V_{in})^2\,dt\right\}\\&= (A\bullet\omega_{in})^2\bullet\varepsilon(\tau_{jitter}^2)\frac{1}{T_{in}}\int_0^{T_{in}}[\cos(\omega_{in}t)]^2\,dt\end{aligned} \tag{3.12}$$

where, $\varepsilon\{\}$ is expectation (probability) operator and $\varepsilon(\tau_{jitter})$ denotes the probability distribution of random jitter (τ_{jitter}). This probability distribution function can be Gaussian distribution or any other distribution. If the variance of $\varepsilon(\tau_{jitter})$ is σ_τ then the average noise power(σ_{jitter}^2) becomes,

$$\sigma^2_{jitter} = \frac{(A \bullet \omega_{in})^2}{2} \bullet \varepsilon\left(\tau^2_{jitter}\right)$$
$$= \frac{(A \bullet \omega_{in})^2}{2} \bullet \sigma^2_\tau \tag{3.13}$$

The SNR due to the clock jitter is,

$$SNR_{jitter} = \frac{A^2/2}{\sigma^2_{jitter}} = \frac{1}{(\omega_{in}\sigma_\tau)^2} = \frac{1}{(2\pi f_{in}\sigma_\tau)^2} \tag{3.14}$$

Equation 3.14 shows the SNR_{jitter} does not depend on the input amplitude. For a given τ_{jitter} a relation can be established with phase error (θ_p) of the clock by the following equation,

$$\theta_p = 2\pi f_{clock} \cdot \tau_{jitter} \tag{3.15}$$

where f_{clock} is the input clock frequency. Therefore the relation between rms phase error of the input clock (σ_p) and the σ_τ cab be expressed as follows,

$$\sigma_p = 2\pi f_{clock} \cdot \sigma_\tau \tag{3.16}$$

Thus the SNR_{jitter} in terms of the phase error of the input clock can be expressed as,

$$SNR_{jitter} = \frac{1}{\left(\dfrac{f_{in}}{f_{clock}}\sigma_p\right)^2} \tag{3.17}$$

or in decibels,

$$SNR_{jitter} = \left[20\log\left(\frac{f_{clock}}{f_{in}}\right) - 20\log(\sigma_p)\right]dB \tag{3.18}$$

In the above equation a simple relation between the signal to noise ratio degradation due to the phase error of the sampling clock is presented. For an 8 bit of accuracy the required SNR is 50dB. With 3GHz of input signal sampled at 10GS/s the required rms time jitter of the input sampled clock is 170fS. This signal can be generated with the low-phase noise sinusoidal sources (e.g. Agilent E8257D) and subsequently converting the sinusoidal into clock signal by the means of an on-chip limiting amplifier. In the present designs, the limiting amplifiers are implemented with cascaded three stage differential amplifiers.

3. 4. 2. 5. Design optimization of the SEF

The error sources discussed in the earlier sections can be reduced in different fashion. Although the noise degradation due the aperture jitter cannot be improved by the SEF. The total contribution of error charge due to the finite aperture time and the pedestal error can be expressed as,

$$Q_{error} = \int_{t_0}^{t_0+\tau_a} \frac{V_A(t) - V_{out}(t)}{Z_1(t)} dt + \Delta \tau_a \cdot I_1 + C_{bc2} \cdot V_{CLK} + \eta \cdot \int_{-V_{out}}^{V_{out}} C_{be1} dV \qquad (3.19)$$

The effect of the finite aperture time can be reduced by the reducing the rise and fall time of the sampling clock and increasing the bias current I_1 of the SEF. Increasing the size of the C_H can reduce the pedestal error due to the base-emitter capacitance of the transistor Q1. But this may reduce the bandwidth of the SEF and introduce higher harmonics distortion. For a given load resistance of the input buffer (R1 in Fig 3.12) the main optimization is done for the bias current I_1 and the hold capacitance C_H.
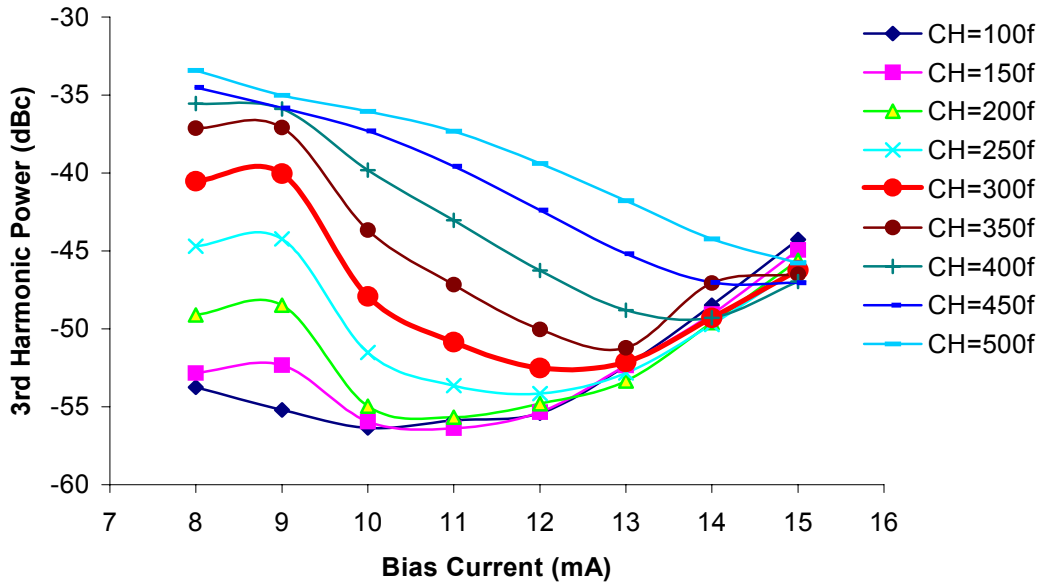
Fig. 3. 16. 3rd harmonic power at SEF of different bias currents and hold capacitances

In Fig. 3. 16 the difference between the fundamental and 3rd order harmonic power is presented for varying bias current (I1) and load capacitance (C_H). In this plot a 3GHz, 2Vpp differential signal is used as the input, which is sampled at 10GHz. For the differential design the even order harmonics are heavily suppressed. Thus for the THD estimation only odd order harmonics are to be considered. Moreover if the 5th and higher order harmonics show much lower amplitude than the 3rd order harmonic, then the difference between the fundamental and the 3rd order harmonic can be approximated as the THD.

With lower bias current the SEF shows higher 3rd order harmonic for a given hold capacitance. This is because of the fact that, in the hold mode the node voltage of A (See Fig. 3. 13) is not sufficient for turning off the base emitter diode of Q1. While with higher bias current the node voltage of A in the hold mode is pulled down to such a lower value that the input transistors of the input buffer go into saturation. As a result the input buffer takes more

time to follow the input signal in the track mode. For high frequency operation the track time is not sufficient and introduces higher error.

The lower hold capacitance shows better linearity because of the faster transition from the track to hold mode. But it produces higher pedestal error droop rate in the hold mode. A compromise can be found for C_H=300fF. It shows THD less than 50dBc (which is consistent with the 8-bit of accuracy) for the bias current I1=12mA. This bias current dictates the emitter size of the transistors used in the design.

3. 4. 3. Output Buffer

The output buffer is a unity gain buffer, which isolates the hold capacitance from the external load. In Fig. 3. 17 two different output buffers are presented. An emitter follower is used in the simple output buffer (see Fig. 3. 17a) to interface the hold capacitor (C_H) and the test buffer. Due to the base current of the bipolar input devices the output buffer causes droop in the held voltage. In general base current compensation technique is used to mitigate this problem [23]. In the output buffer shown if Fig. 3.17a no base current compensation technique is used, so the single-ended droop rate will be comparably high. But the high update rate of 10GS/s reduces the droop rate effect. Further, the symmetry of the circuit provides droop compensation, resulting in an acceptable droop error. A simple emitter follower with a resistive divider is used as a test buffer. The resistors R5 and R6 have values of 450 and 50 Ohms, respectively.

(a)   (b)
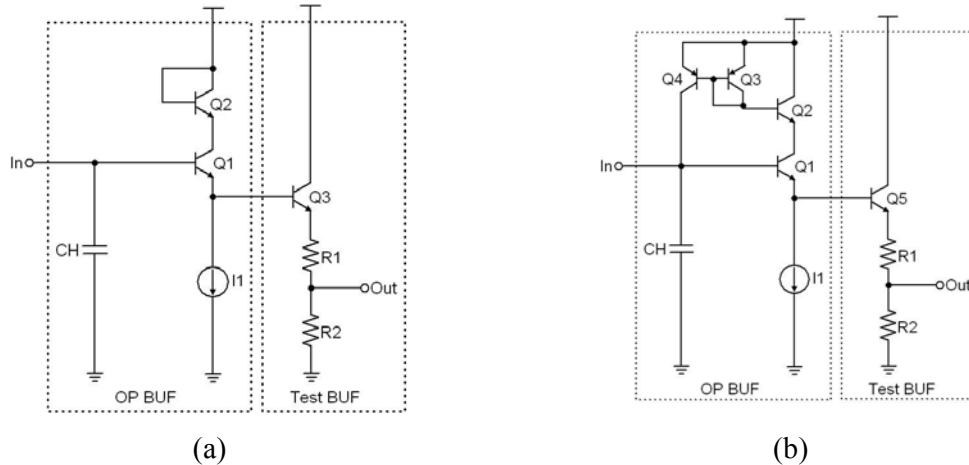
Fig. 3. 17. (a) Simple output buffer (b) Output buffer with base current compensation

An improved output buffer is proposed in Fig. 3. 17b. A pnp current mirror is used to improve the droop rate by compensating the base current of Q1 with the replica of base

current of Q2. This current mirror can track the base current much faster than the conventional pMOS mirror [23]. As a result, the base current compensation is more effective over a wider range of the signal. The test buffer has the similar configuration of Fig. 3. 17a.

3. 4. 4. Implementation of Full THA

Two different THAs are implemented by combing the sub-blocks described above. In the first variant only npn transistors are used. The simplified schematic of npn THA is presented in Fig. 3. 18 [29]. For the first time in the context of open loop THA, a cascode amplifier is used as an input buffer to enhance the input range. A simple SEF is used to accomplish the sampling operation and an emitter follower is used as the output. Three different power supplies are used for the input buffer, SEF and the test buffer. This circuit is optimized for a 2Vpp differential sinusoidal and the input bandwidth is 3GHz with a sampling rate of 10GHz.
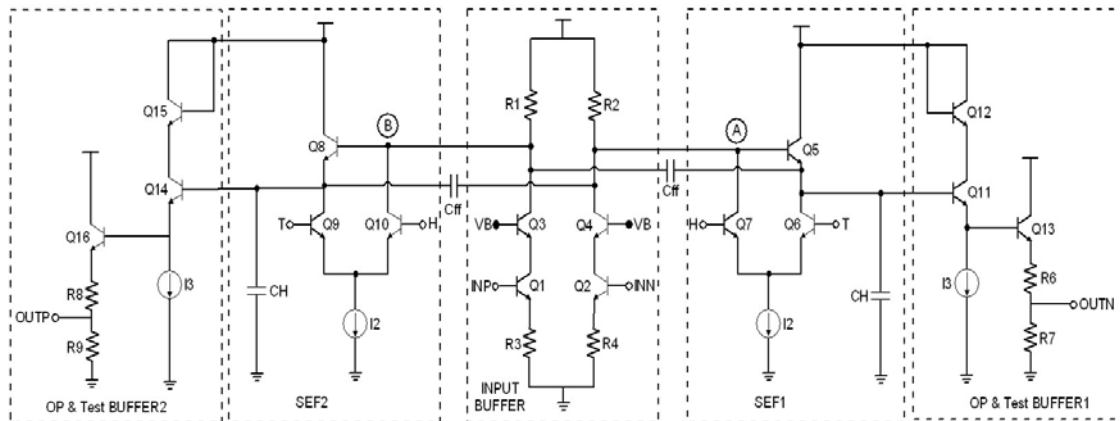
Fig. 3. 18. Simplified schematic of npn THA

The second THA is implemented with the novel complementary npn pnp transistors [28] (see Fig. 3. 19). From the input to the output only emitter followers are used in the main signal path. This provides well-defined gain, near to unity at the output of the THA. Further it enhances the bandwidth of the full THA. The SEF is same as used in the npn THA. At the output buffer base current compensation technique is used to improve the droop rate in the hold mode. In general BiCMOS process the base current compensation is implemented by the pMOS current mirrors [23]. But in the high-speed applications this compensation process is not precise due to the slow pMOS devices. In the current application the technology provides

43

pnp devices which make it possible to implement the base current compensation loop with the pnp devices.

Both of the THAs have pseudo differential architecture. Therefore special care has been taken for the matching among the active devices. Transistors with larger emitter area have been used in the main signal paths. 50Ω microstrip transmission lines are used to connect the inputs and outputs of the core circuitry to the external pads.
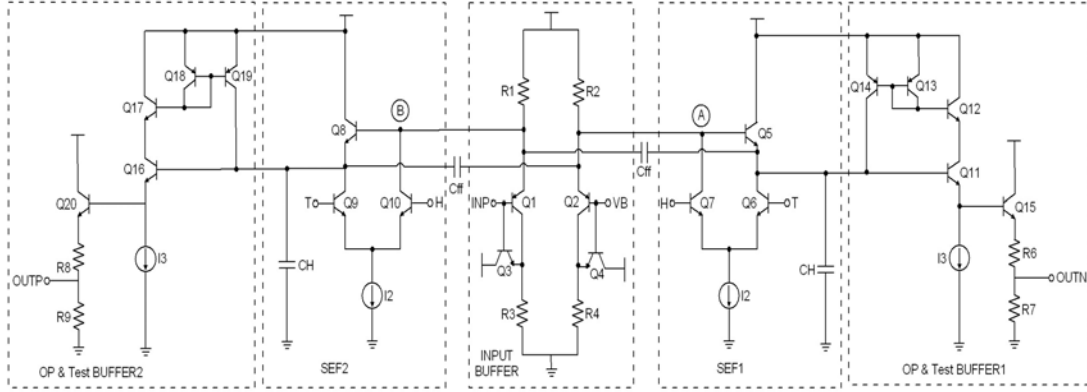
Fig. 3. 19. Simplified schematic of npn pnp THA

3. 5. Double Sampling THA

The conventional THA architectures presented in the previous section have been implemented with so called single sampling technique. In this method, ideally half of the total time period is dedicated to track the input signal and rest of the time is spent to hold a valid sampled voltage. Therefore almost 50% of the clock cycle the THA has an invalid output. This reduces the time available for further processing of the sampled value (e. g. the quantization process). For instance, with 100ps sampling period, only 50ps would be available for processing after the THA, which imposes strong constraints on the following stages, e.g. comparators.

One of the well-known methods to enhance the sampling rate is known as double sampling technique. In this method instead of a single sampling capacitor a pair of capacitors is used and two parallel sampling switches are controlled with 180˚ phase-shifted clocks. When one switch tracks the input signal the other one works in hold mode. At the output of the THA, only the hold mode outputs of the sampling switches are combined alternatively. This provides much more time for further processing for the subsequent stages.

Although this double sampling technique is popular in closed loop THA design [30], but it has rarely been used in open loop architectures. An open loop THA architecture using the

44

double sampling technique is presented in [31]. The THA is implemented with three main blocks. An input multiplexer selects either of the two parallel sampling modules alternatively. Then the core-sampling module stores the sampled voltage across the hold capacitors and an output multiplexer is used to combine the hold mode outputs.

The main drawback of the architecture arises from the clock-skew [30]. Unequal zero crossing intervals among the differential clocks appear as the clock jitter and deteriorate the resolution of the THA. In this present work an attempt has been made to overcome the aforementioned problem by developing a new open loop time skew insensitive SEF [32].

A simplified block diagram of the proposed pseudo-differential THA architecture is shown in Fig. 3. 20. In this figure only one of the pseudo-differential paths (Vin→OP) is shown. The other path (Vin_B→OP_B) uses another identical block. Unlike the conventional open loop THA the input buffer is connected with two parallel SEFs. The clock timing of the SEFs is explained in the later section. The outputs of the SEFs are combined by using an analog multiplexer driven by another clock CLK3 which can be derived from either CLK2 or CLK2B. The analog multiplexer is optional in some applications particularly when two parallel quantizers are used. The output of the parallel switches can be directly connected to the input of the quantizers.
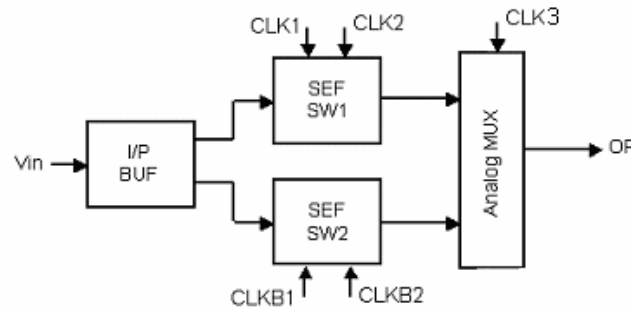
Fig. 3. 20. Block diagram of proposed pseudo-differential double sampling open-loop THA

3. 5. 1. Input Buffer

Fig. 3.20 shows that each of the pseudo-differential paths of the double sampling THA has only one input buffer. This input buffer is shared by the parallel SEFs (SEF SW1 and SEF SW2) which are eventually clocked in time interleaved fashion. As a result of it the input buffer is always active and following the input signal. If the switching mechanism of the THA shown in Fig. 3.18 is reconsidered then it can be seen that during the hold mode the an excess current is drawn form the load resistor. As a result of it the output of the input

buffer is destroyed during the hold mode and it can no longer be used by the other SEF. On the other hand the THA architecture described section 3. 3. 2 is particularly suitable for the time multiplexing the SEFs. In this section the modified version of the input buffer is described which has two output branches.
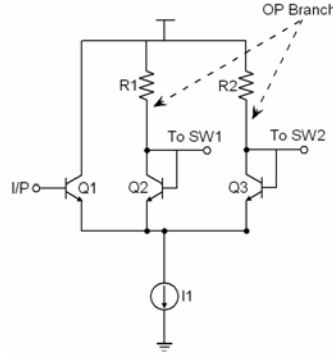
Fig. 3. 21. Input buffer of double sampled THA

Schematic of the proposed input buffer for the double sampling THA is shown in Fig. 3. 21. This unity feedback input buffer has the similar structure as explained in section 3. 3. 2. The main difference can be found in the output branch. Unlike [33] two output branches are used here which are connected to a pair of parallel SEFs (SW1 & SW2). For any point of time either of the output branches is conducting and the other branch is off. This is accomplished by the SEF switching. When an SEF (SW1 or SW2 of Fig. 3.20) is in track mode the corresponding output branch of the input buffer is following the input signal. Whereas in the hold mode the base-emitter voltage of the same output branch transistor (Q2 or Q3) is pulled down to zero and restricts it to follow the input analog signal.

Generally pMOS current sources are used as a load for the input buffer [23]. But with the increasing input frequency the output impedance of the pMOS current source is reduced and the gain of the input buffer drops down. For a large input bandwidth this causes a frequency dependent input buffer gain. To overcome this problem, here a resistor is used as load instead of the pMOS (Fig. 6). This provides much higher bandwidth. It may reduce the gain of the input buffer, but this can be compensated later in the output stage of the THA.

3. 5. 2. Skew Insensitive Double sampling SEF

In a double sampling THA architecture two parallel sampling switches are used. These switches are controlled by perfectly inverted clocks (T & H) as shown in Fig. 3. 22 (ideal clock). In the context of SEF, each of the zero crossings of the differential clocks T and H precisely defines the sampling instant of respective sampling switch. In reality it is very

46

difficult to generate such a perfectly differential clock. Due to the device mismatch and unequal parasitic capacitance, the duration between two consecutive zero crossings can have some error (as shown in Fig. 3. 22, Real clock). This error comes as a sampling jitter and deteriorates the performance of the full THA particularly when the sampling rate is in gigahertz range.
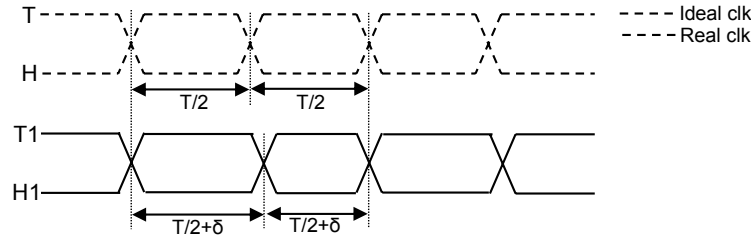
Fig. 3. 22. Clock timing skew

To overcome this problem a new improved time skew insensitive SEF schematic is proposed [32] in Fig. 3. 24. The basic switching principle is similar to the SEF explained in the section3.3.2. Unlike the SEF presented in [3], two stacked pairs of differential current switches are used here. The upper differential current switch consists of Q2 and Q3 and it is controlled by the differential clock CLK1 and CKLB1. The lower differential current switch (Q8 & Q9) is controlled by the differential clock signal CLK2 and CLKB2. CLK2 has half of the frequency of CLK1 (Fig. 3. 24). The duty cycle of CLK1 can be varied to maximize the tracking time of a particular switch. The only constrain is that, the falling edge of CLK1 must appear earlier than the zero crossing of CLK2 and CLKB2.

As explained in the section 3.3.2, a voltage shifted version of the sampled signal is feedback the transistor Qclp in the hold mode to reduce the feedthrough. This requires an additional SEF to accomplice the feedback. In this current design a DC voltage (V_{DC}) is connected to the input of the Qclp, which is sufficient to turn off Q1, even with the minimum input level.

The schematic of one of two pseudo differential paths of the proposed double sampling THA is presented in Fig. 3. 25. Either of the SEF1 or SEF2 goes into track mode in an interleaved fashion. Assuming CLK2 is high the SEF2 goes into track mode when CLKB1 goes high. In this mode Q1 and Q2 appear as the unity feedback input buffer. Transistor Q9 of SEF2 acts as the emitter follower and tracks the input voltage. When the CLKB1 transits from high to low and CLK1 goes low to high, at the zero crossing point track mode ends. On the other hand with CLKB1 being in high state, SEF1 operates in the hold mode. In this mode

47
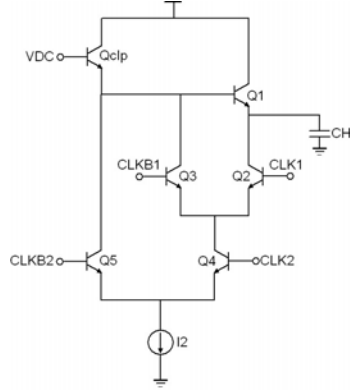
Fig. 3. 23 Schematic of double sampled SEF
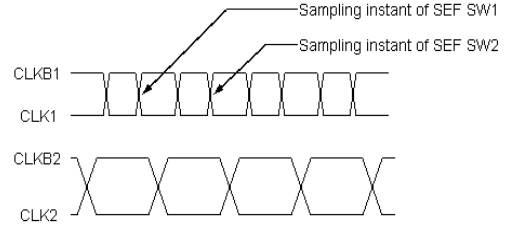
Fig. 3. 24. Timing diagram of double sampling SEF

Qclp1 appears as an emitter follower. The level shifted sampled voltage turns off Q4. Thus the sampled voltage is stored across the sampling capacitor C_H. Additionally collector-emitter voltage of Q3 is pulled down to zero [23], which restricts Q2 from following the input voltage.
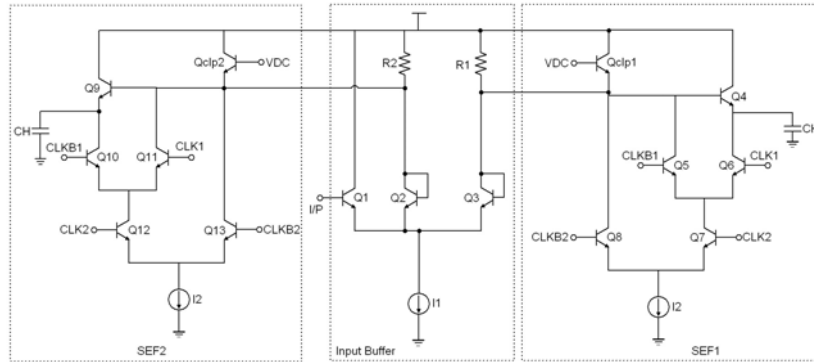
Fig. 3. 25. Schematic of a pseudo differential path of the core double sampling THA

3. 5. 3. Analog Multiplexer

A possible schematic of the analog multiplexer is shown in Fig. 3. 26. It has a structure similar to an ECL D-latch but without feedback. The main operation of this circuitry is to select either of the input pairs (INP, INN or INP1, INN1) alternatively and pass it to the output of the THA. Transistor pairs Q1, Q2 and Q3, Q4 work as input emitter degenerated differential amplifier. Whereas the lower differential pair Q5 and Q6 operates as a current switch.

Still, this solution has the limitation that, when either of the input differential pairs is off, the parasitic base-collector capacitance of the inactive input transistors couples unwanted

48

signal to the output node. As a result of it the hold mode signal gets distorted and the effective resolution of the THA is reduced. This can be compensated by the well-known measures to reduce hold-mode feedtrough, e.g. increase the size of the hold capacitor C_H, minimize the size of the transistors in the differential pairs, and add some circuitry to force the base nodes of the inactive pair to some fixed potential.
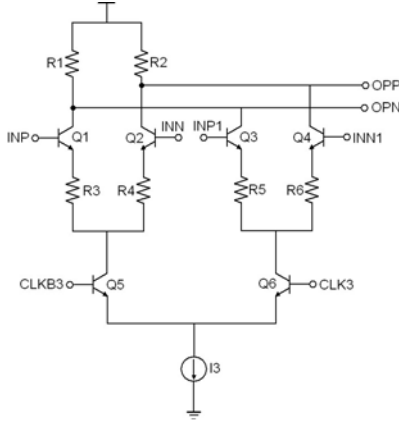
Fig. 3. 26. Schematic of Analog multiplexer circuit

3. 5. 4. Preliminary simulation results

The core double sampled THA (as shown in Fig. 3. 25) is implemented to verify the principle of time skew insensitive SEF. The circuit is optimized for 2GHz of input bandwidth with sampling rate of 10GHz. A 4.5V power supply is used for both input buffer and sampling switches which results in the power consumption of 280mW. The pseudo differential outputs of two parallel paths are shown in Fig. 3. 27 for 1GHz 1Vpp differential input sinusoidal sampled at 10GS/s. Output multiplexer is not included in the simulation. Although an ideal analog multiplexer is used to combine the outputs to estimate the error of the core THA. The transient response of the combined parallel pseudo-differential outputs for the same 1GHz input sampled with a 10GHz clock is plotted in Fig. 3.28. It shows a differential droop rate <1mV/nS.

The accuracy of the THA is simulated in frequency domain. THD is approximated as the difference between the fundamental and the 3rd order harmonic as all other odd harmonics have much lower amplitude. A plot of fundamental and 3rd order harmonic power is presented in Fig. 3. 29 for 1Vpp differential signal sampled at 10GHz. The core THA shows almost a flat 3rd order harmonic for the bandwidth of 2GHz. If the input frequency is increased beyond one-fourth of the sampling frequency (Fs) then the first order inter-

modulation product stem from CLK2 (0.5Fs-Fin) appears within the input bandwidth and dictates the accuracy of the THA.
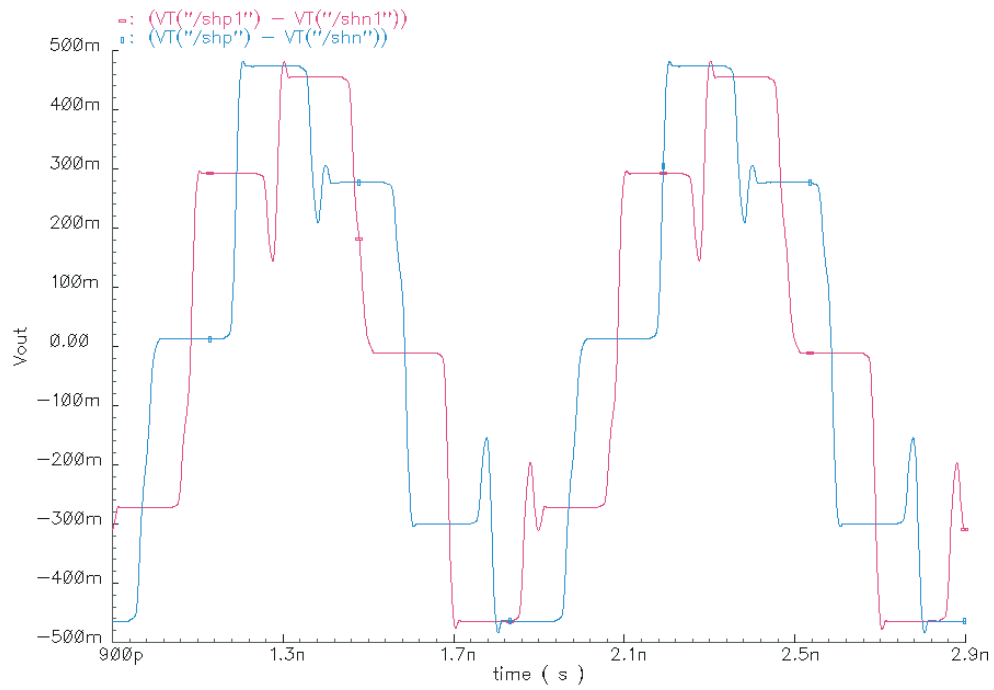
Fig. 3. 27. Transient response of parallel pseudo differential output
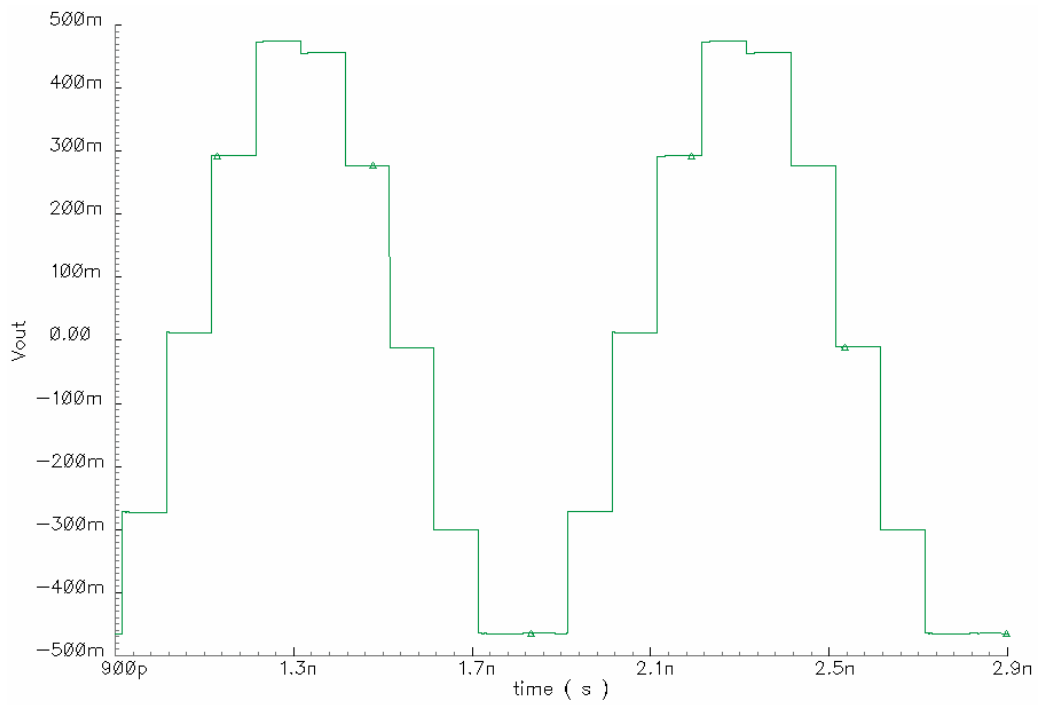
Fig. 3. 28 Combined outputs of the parallel paths of double sampling THA
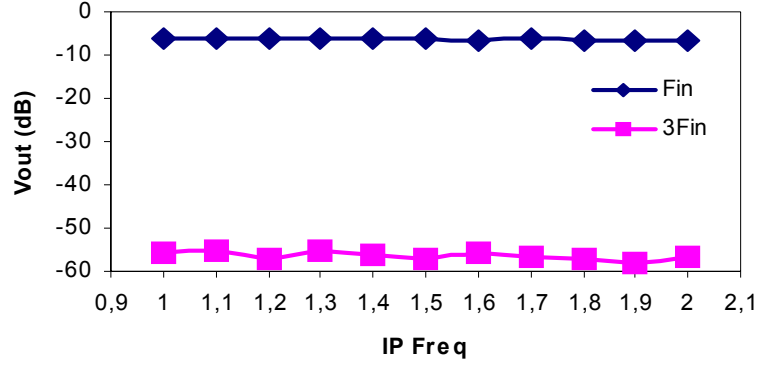
Fig. 3. 29. Spectral components of the double sampling THA

In Fig. 3. 30 the output spectrum of the THA for 2GHz 1Vpp sinusoidal is presented the. It shows THD of -47.94dBc, which corresponds to 7.95 bits of accuracy. A summarized performance of the THA is presented in Table 3.1.
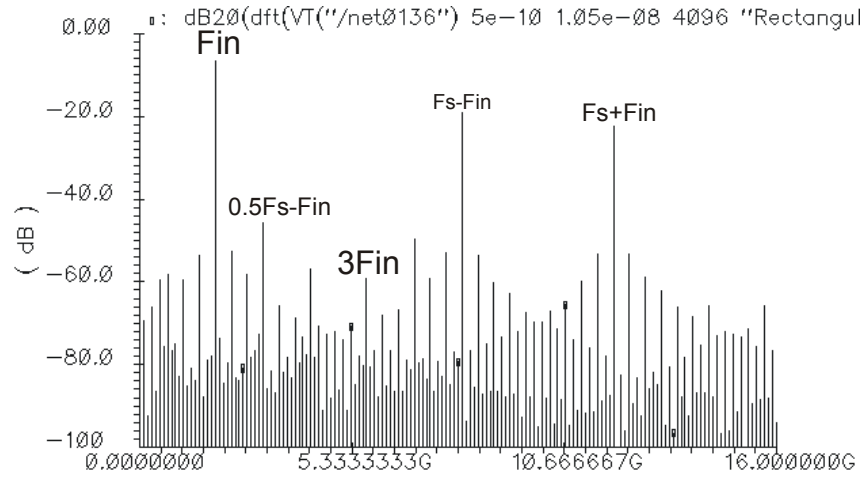
Fig. 3. 30. Output spectrum of double sampled THA

Table. 3.1. Simulated performance summary of double sampled THA

Process	0.25µm 190GHz SiGe BiCMOS
Input range	1Vpp differential
Sampling rate	10GHz
Effective resolution bandwidth	2GHz
THD@F_{in}=2GHz,F_s=10GHz	-47.94dBc
ENOB @ F_{in}=3GHz,F_s=10GHz	7.95bits
Supply voltage	4.5V
Power dissipation of the core	280mW

3. 6. Experimental Results of implemented THAs

The npn THA presented in Fig. 3. 18 has been implemented in a commercially available 0.25 µm 200 GHz BiCMOS technology [34]. The chip micrograph of the THA is shown in Fig. 3. 31. The core area is 0.27mm^2 and the total chip area is 0.97mm^2. The input buffer operates with 5.5 V and rest of the circuit operates with 4.5 V. It dissipates 800 mW of power, including an on chip clock driver, which dissipates 315 mW and was not optimized for low power.

The test setup for the THA is shown in Fig. 3. 32. The chip is wire-bonded on a ceramic board for characterization. A differential signal at the input is provided using an 180º hybrid together with adjustable phase-tuners to produce perfect 180º phase shift at the input pads of the THA. Two pseudo differential outputs of the THA are having different delays due to unequal bond wire length, path delays of the board and coaxial cables. Although for differential measurement these delays have to be equal. But unfortunately another pair of phase-tuners was not available to compensate these off-chip delays. So a single-ended output is used to characterize the THA. A single-ended external clock is used for sampling. The output buffer of the THA uses a resistor divider consisting of 450 and 50 ohms in series to match with external 50 ohms load, causing an attenuation factor of 19.
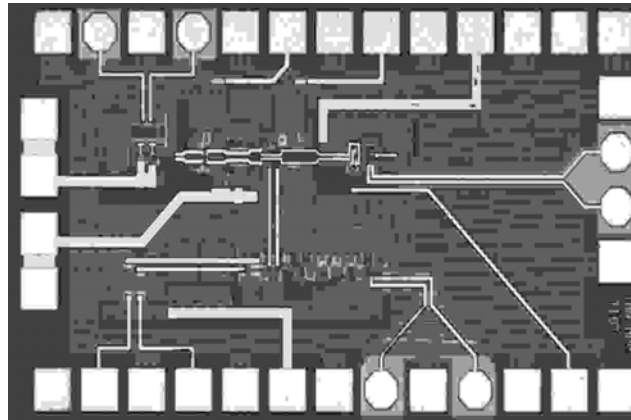
Fig. 3. 31. Chip micrograph of npn THA

Fig. 3. 32 shows the measured single-ended spectrum of the THA at 2 Vpp differential, 3 GHz input (Fin) and 10 GHz sampling frequency (Fs). The second order harmonic is not suppressed in single-ended mode, causing unrealistically high amplitude. The 3rd order harmonic is at -47.22 dBc.
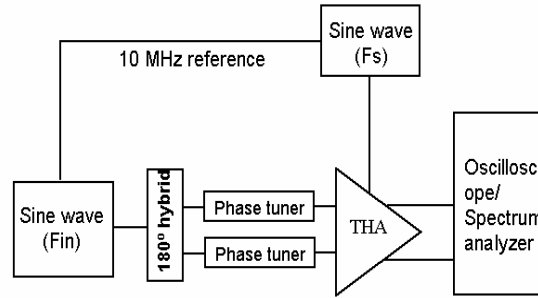
Fig. 3. 32. Test setup for characterizing the THA

In Fig. 3.34, the measured spectral components of Fin (3 GHz) up to the 3rd order harmonics are plotted for different input level at a sampling rate of 10 GHz. It shows that both of the pseudo-differential outputs of the THA are having almost the same amplitude for Fin and its harmonics. Hence, it can be concluded, for differential output the even order harmonic suppression would be sufficiently high to neglect the even order harmonics in the total harmonic distortion (THD) estimation. So the difference between the fundamental and 3rd order harmonic can be approximated as the THD. For 3 GHz 2 Vpp differential signal it is -47.22 dBc, which corresponds to 7.58 ENOB.
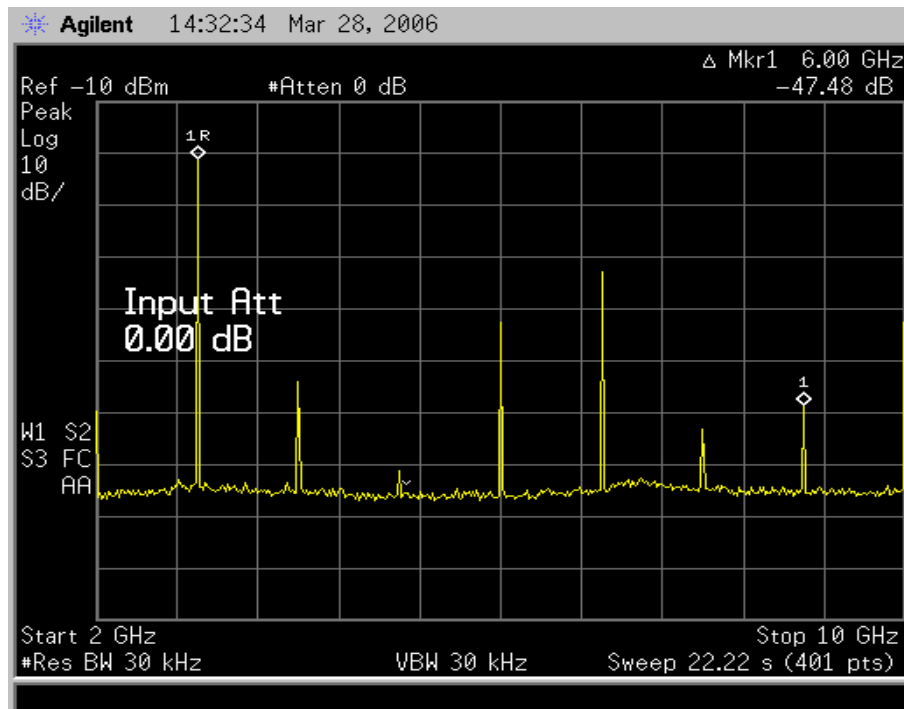
Fig. 3.33. Measured single-ended frequency spectrum of the THA

53
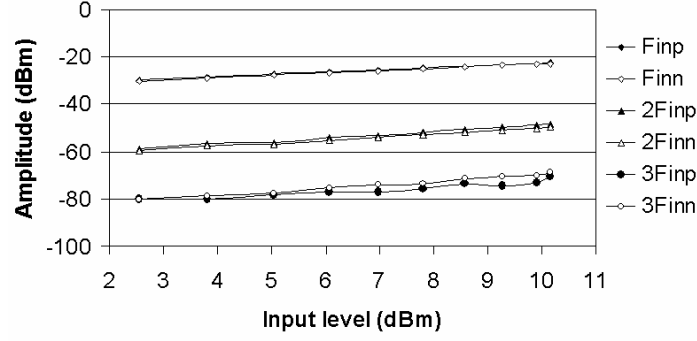
Fig. 3. 34. Measured spectral components of pseudo-differential outputs

Fig. 3.35 shows a time-domain waveform of the THA for 2 GHz, 2 Vpp differential signal input sampled at 12GHz which is 20% higher than the sampling rate used for frequency domain characterization. Two pseudo-differential outputs are separately shown. The oscilloscope 'Math' function is used to plot the difference between these two pseudo-differential outputs. Due to the off-chip delay mismatch droop errors of the pseudo-differential paths are not totally compensated. As a result of it, the differential signal output shows higher droop rate.
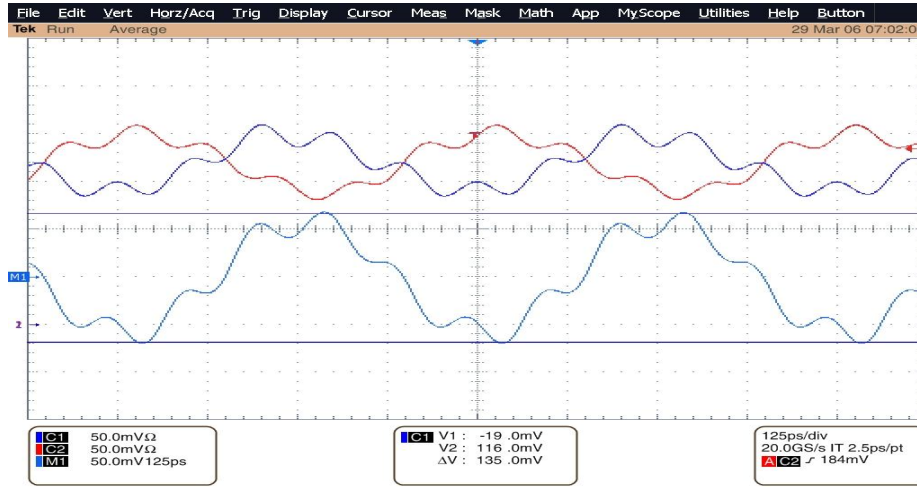
Fig. 3. 35. Measured output waveform at 12 Gs/s with 2 GHz 2 Vpp input

A simplified schematic of the THA is presented in Fig. 3. 19. This THA has been implemented using IHP's 0.25μm BiCMOS complementary HBT technology with f_T=185GHz/90GHz for the npn/pnp HBT [35]. The layout of the THA is shown in Fig. 3. 36. The core area is 0.34mm^2 and the total chip area is 1.65mm^2. It consumes 587.5mW of power from 5.0V power supply.
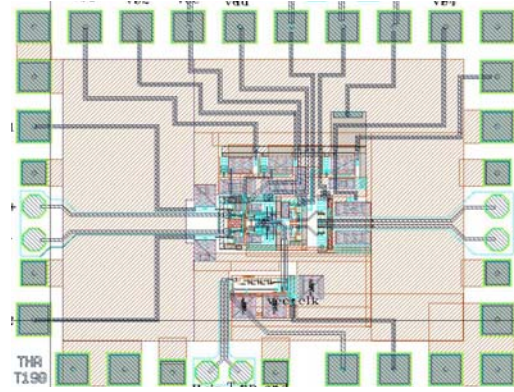
Fig. 3. 36. Layout of npn pnp THA

The chip is tested on-wafer with a 40GHz probe station. For critical inputs and outputs 40GHz coaxial cables were used. A similar test setup as shown in Fig. 3. 32 is used to characterize the THA. The transient response of the THA with 1GHz 1Vpp differential sinusoidal input (Fin) sampled at 10GHz is shown in Fig. 3. 37.
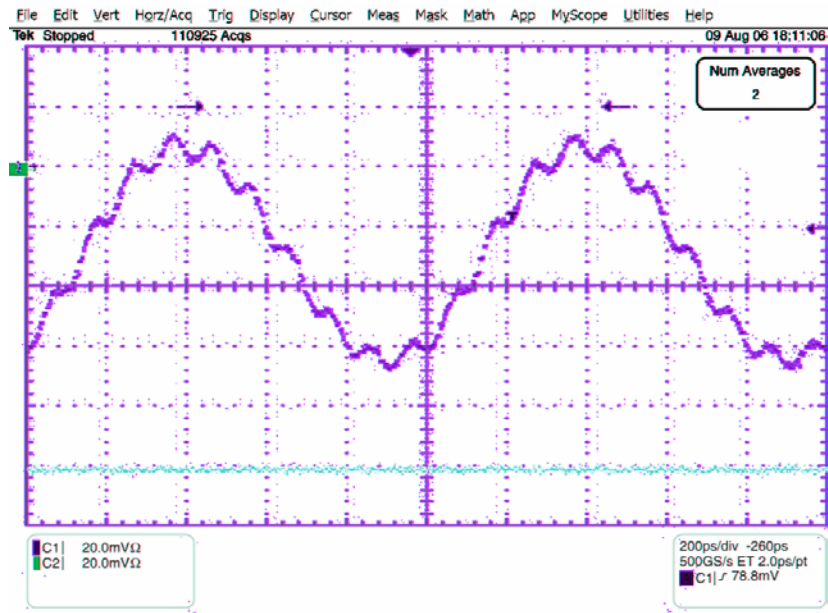
Fig. 3. 37. Transient response of npn pnp THA for Fin=1GHz @10Gs/S

With the same input and sampling rate the THA shows 3rd order harmonic at -38.92dBc. The single ended spectrum of the THA is shown in Fig. 3.38.This corresponds to 6. 2 ENOB. The THA was simulated for 8-bit of accuracy with 2Vpp differential input. But unfortunately some errors were found in the pnp compact model (VBIC), which can be identified as the main reason for the discrepancy between the simulation and the measurement results. The

summarized performance of the npn pnp THA as well as the npn THA are presented in Table 3. 2.
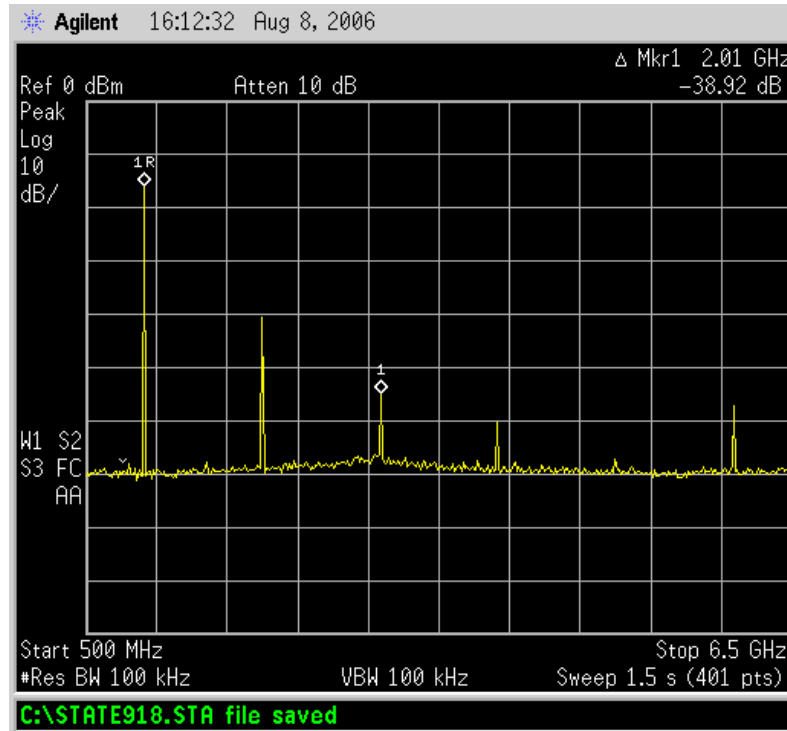
Fig. 3. 38. Single output spectrum of npn pnp THA for Fin=1GHz and Fs=10GHz

Table. 3. 2. Performance summery npn and npn pnp THAs

Parameter	npn THA	npn pnp THA
Process	0.25µm 190GHz SiGe BiCMOS	0.25µm SiGe BiCMOS with complementary HBTs
Input range	2Vpp differential	1Vpp differential
Sampling rate	10GHz	10GHz
Effective resolution bandwidth	3GHz	1GHz
THD	-47.22dBc	-38.92dBc
ENOB	7.58bits	6.2 bits
Supply voltage	5.5/4.5V	5.0V
Total power dissipation	800mW	587.5mW
Die area with pads	0.97mm^2	1.65mm^2

56

Table 3. 3 compares published high-speed Si/SiGe THAs with the present works. The best performance in terms of sampling frequency can be found in [38]. This THA is implemented 0.18 μm BiCMOS technology. A complex distributed sampling technique is used, where three separate THA modules are used in pipeline. The accuracy of the circuit strongly depends upon the on-chip delay lines. Practically one THA module can sample at 16.66GHz, with 6bit of accuracy. By using the conventional open loop architectures the best results in terms of sampling speed and resolution is reported in [27]. It achieves 8.0 ENOB for 12.1GHz of sampling rate. This work outperforms [27] in terms of input range and input bandwidth. It is the only THA known to the authors which achieves >5 bits effective resolution for a 2 Vpp 3 GHz input signal.

Table 3. 3 Comparison with published Si/SiGe high speed THAs in SiGe technology

Ref. No.	Fs [GHz]	Fin [GHz]	Input [Vpp]	ENOB [Bits]	Supply [V]	P_{diss} [mW.]	Process/f_T
[24]	10.0	1.0	1.0	6.8	3.3	70.0	SiGe/200
[27]	12.1	1.5	1.0	8.0	3.5	700.0	SiGe/200
[36]	1.2	0.6	1.0	8.0	+2.0/-0.5	460.0	Si/25
[37]	18.0	2.0	1.0	5.0	3.5	128.0	SiGe/120
[38]	50	40		6.0	4.0/3.3	640	SiGe
[39]	40	19		4.2	3.6	540	SiGe/160
npn pnp THA	10	1	1.0	6.2	5.0	587.5mW	SiGe BiCMOS complementary HBTs
npn THA	10	3.0	2.0	7.58	5.5/4.5	800.0	SiGe/200

3. 7. Design of High-Speed Comparator

The analog to digital conversion process can be divided into two main operations. The sampling process is accomplished by the front end THA. Then the time discrete analog signal is approximated to the predefined reference voltage by a quantizer. A comparator comes as a building block for the quantizer design. The accuracy of the quantizer mainly depends upon the accuracy of the comparator.

In the previous sections the design of three multi-GHz THAs has been presented. The npn THA shows highest accuracy of 7.83ENOB. By using this front-end THA an 8-bit ADC can be built. In chapter 2 it has been shown that the flash or folding-interpolating architecture are the most suitable for high-speed applications. But 8-bit flash ADC would not be efficient in terms of power and area. The speed, power tread-off can be found in folding-interpolating architecture. In section 2.5.3 a description of folding architecture has been presented. It has been assumed that 2-bit of coarse quantization is used rest of the 6-bits are resolved by the folding-interpolating stage. For this 6-bit folding-interpolating sub-ADC a 6-bit accurate comparator comes as basic building block. In the following section the design of a open-loop high-speed comparator is presented to fulfill the speed and accuracy requirements of the 6-bit sub-ADC.

The basic comparator architecture is consists of two main parts; the preamplifier and the regenerative latch stage. In MHz sampling rate regime CMOS comparators are preferred over the bipolar counterparts because in general those comparators do not have static power dissipation. But in multi-GHz range sampling rate the bipolar comparators come as the obvious choice. High sampling rate is achieved with high power dissipation [40], [41].
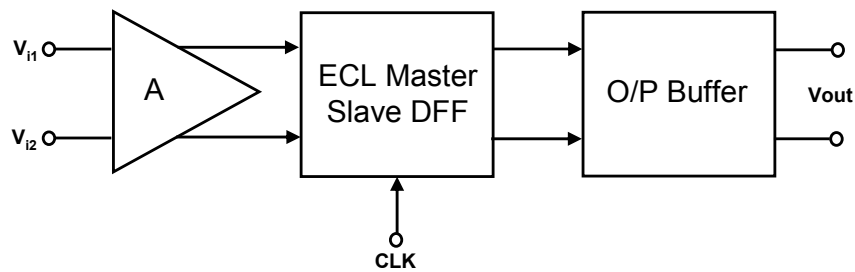
Fig. 3. 39. Block diagram of high-speed comparator

A general block diagram of high-speed comparator is presented in Fig. 3.39 [4]. The comparator has two main blocks; the preamplifier and a positive feedback latch. An output buffer is introduced to drive the external 50Ω load. In the high speed comparator design a differential amplifier is used as a preamplifier. The preamplifier substantially reduces the kickback noise [4]. Secondly it provides additional DC gain. This additional DC gain in turn reduces the input referred offset of the comparator by reducing the contribution of the latch input referred offset. In Fig. 3.40 the simplified schematic diagram of the preamplifier is presented. It is implemented with a single stage differential amplifier. Pair of emitter followers is used as the input buffer for the preamplifier. The emitter followers input impedance is directly matched with the 50Ω through the fixed bias resistances R1 and R2. At the output of the preamplifier another pair of emitter followers is used to match the output

58

DC level of the preamplifier with the input DC level of the master slave DFF. These emitter followers are also very instrumental to isolate the master latch from the preamplifier.
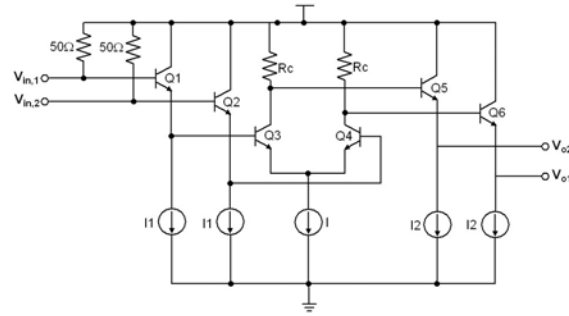
Fig. 3. 40. Simplified schematic of the preamplifier

The regenerative latch of the comparator is implemented with the conventional ECL master slave DFF (MSDFF). A simplified block diagram of this ECL MSDFF is presented in Fig. 3.41. The MSDFF is consists of two identical ECL D-latches. The input differential signals (D, DB) are connected with the master latch and the slave latch provides the differential output (Q, QB). These cascaded master and slave ECL D-latches are controlled by a differential input clock (CLK, CLKB). The master and slave latches work in the time interleaved fashion. This is accomplished by twisting the differential clock at the clock input of the slave D-latch. The schematic of a commonly used D latch is presented in Fig. 3.42.
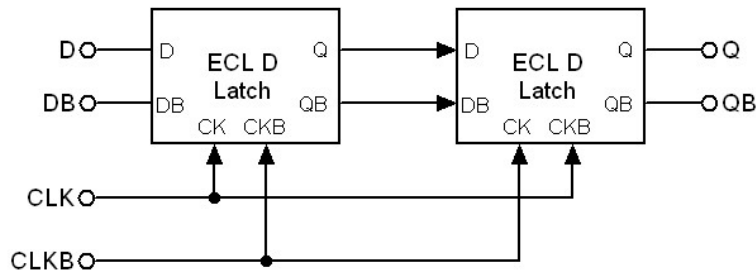
Fig. 3.41. Block diagram of ECL master slave DFF

The latch has two phases. In the first phase the differential clock signal CK goes higher than the CKB signal and the tail current (I1) is switched through the transistor Q5. In this mode the input transistors Q1 and Q2 tracks the input signal. In the next phase the CKB signal goes higher with respect to CK and the tail current switches to emitter coupled pair Q3 and Q4. In this phase the emitter followers Q7 and Q8 provide the positive feedback and the output levels at the output are latched. The detail description of the ECL D-Latch is presented in Chapter 5.
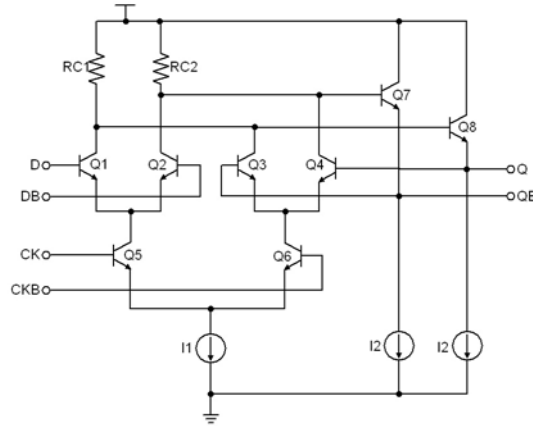
Fig. 3.42. Simplified schematic of D latch

The output buffer is implemented with two stage cascaded differential amplifiers. The final differential stage of the output buffer is has load resistance of 50Ω. To drive the output load this stage has high bias current.

3. 8. Measurement Results of the Comparator
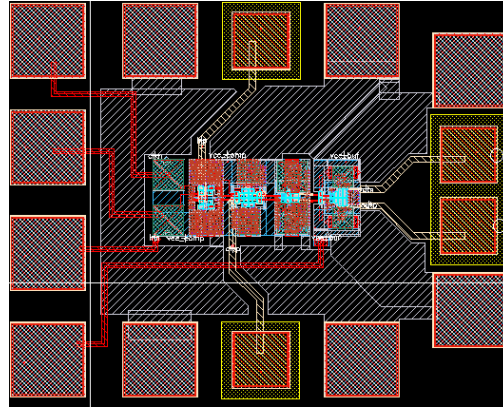
Fig. 3. 43. Layout of 20GHz HBT comparator

The high-speed HBT comparator has been implemented in IHP's $0.25\,\mu m$ $200\,GHz$ BiCMOS technology SG25H1 [34]. The chip layout of the comparator is shown in Fig. 3.43. The core area is $0.07\,mm^2$ and the total chip area is $0.45\,mm^2$. The full chip operates with 5.0V of power supply. Core comparator consumes 70mW of power and the output buffer together with clock buffer consumes 250mW of power. The output buffer was not optimized for power. It was mainly designed drive the measuring instrument (50Ω load). The comparator was tested on-wafer with a 40GHz probe station. For critical inputs and outputs 40GHz coaxial cables were used. The test setup is presented in Fig. 3.44. A low phase noise sinusoidal signal from external signal source was used as input clock. Since the output buffer

of the comparator is matched with the external 50Ω load, it was possible to connect the outputs directly to the sampling oscilloscope through DC blockers. A frequency divided clock signal is used to trigger the sampling oscilloscope. The measurement is done with a single ended input signal. A DC source is used to generate the reference for the comparator. The measurement is done in the time domain.
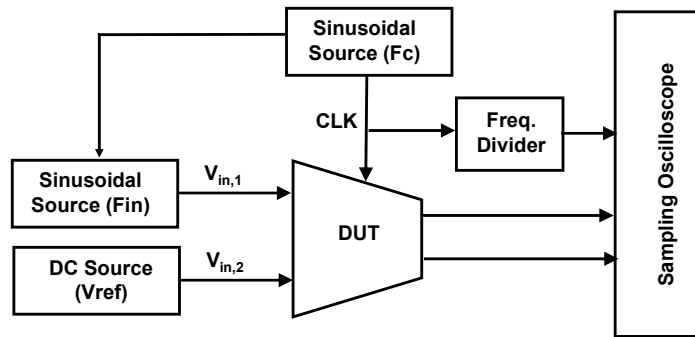
Fig. 3. 44. Test setup for the comparator

For the accuracy measurement the amplitude of the input sinusoidal source is varied and the output waveform of the comparator is observed in the sampling oscilloscope. In Fig. 3.45 a magnified output of the comparator is shown, where a 2GHz 100mVpp sinusoidal is used as the input signal (F_{in}) and the reference voltage (V_{ref}) is at the middle of the input sinusoidal. A 20GHz clock (F_c) is used for this measurement. The rise and fall time of the comparator output is measured to be 15ps which is comparable with the simulated rise
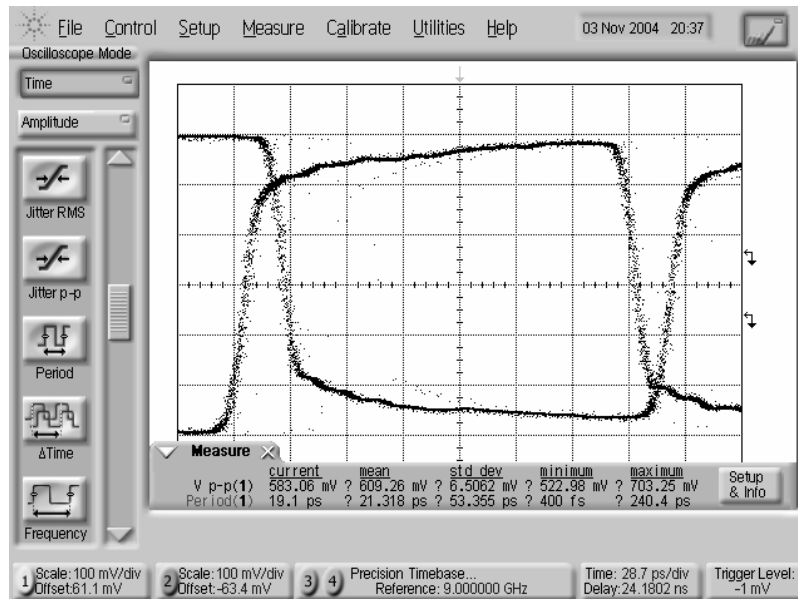
Fig. 3.45. Magnified output waveform of the comparator for 2GHz 100mVpp sinusoidal with 20GHz of clock

and fall time of the output buffer. The differential outputs show 50% duty cycle.

With the decreasing input amplitude of the comparator the output wave from starts to deviate from its symmetrical behavior. The main reason is the input referred offset of the comparator which in turn dictates the resolution of the full comparator. Fig. 3.46 shows the output of the comparator for 25Vpp sinusoidal input with 20GHz of clock rate. The output waveform is already distorted due to the input offset. The measured resolution of the comparator with 2GHz input signal and 20GHz clock is 17.5mV. For a 1V full-scale input it corresponds to 5.8-bit of accuracy. In table 3.4 the summarized measurement results of the comparator is presented.
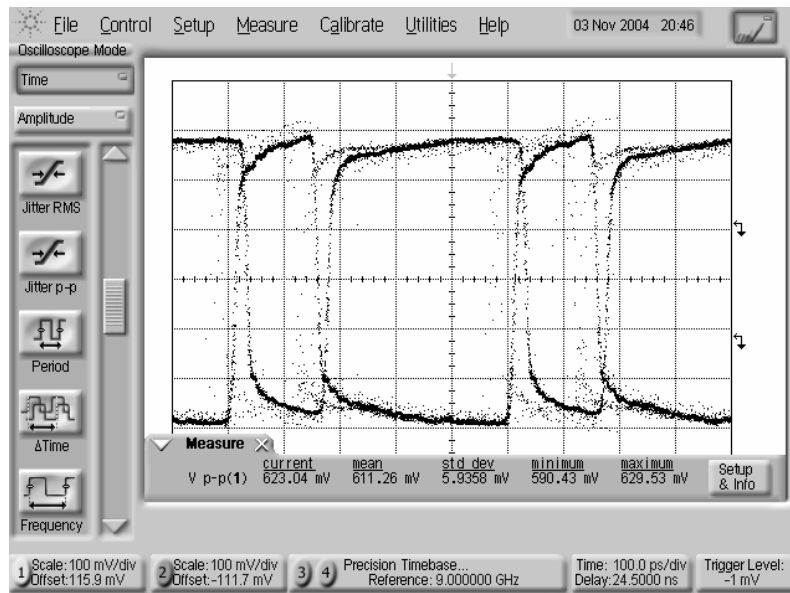
Fig. 3.46. Oput waveform of the comparator for 2GHz 20mVpp sinusoidal with 20GHz of clock

Table 3.4. Summary of measurement results

Process	IHP's 0.25μm SiGe BiCMOS SGC25C
Resolution	5.8 bit
Conversion rate	20 GHz
Input bandwidth	2 GHz
Supply voltage	5.0 V
Core power dissipation	70mW
Die area with pads	0.45mm^2

3. 9. Conclusions

In this chapter an open-loop THA in a 190 GHz SiGe BiCMOS technology is presented. A pseudo-differential npn cascode stage is used as the input buffer, which increases the input voltage swing up to 2 Vpp differential. It achieves 7.58 bits of accuracy at 10 GS/s of sampling rate with 3 GHz of input bandwidth. Compared to the published high-speed THAs, the current work has better performance in terms of input range and bandwidth. At the same 2 Vpp swing, the improvement in ENOB is about three bits.

In the second implementation an emitter follower only THA circuit is presented. An adaptively V_{CE} adjusted npn pnp emitter follower is used as the input buffer to increases the input voltage swing. It achieves 6.2 bits of accuracy at 10 GHz of sampling rate with 1 GHz of input bandwidth.

To increase the sampling rate, a double sampled open loop THA architecture is proposed. The main source of error in this double sampling technique is identified as the time skew in between the parallel sampling switches. To overcome this problem an improved time skew insensitive SEF structure is proposed. To verify the operating principle of the proposed SEF an open loop double sampled THA is implemented which shows 7.5 ENOB for an input bandwidth of 2 GHz at 10 GHz of sampling rate.

As the basic building block of a quantizer an open loop comparator is designed which can be used to build a 8-bit folding interpolating ADC. In this comparator design the continuous time preamplifier is used and the regenerative latch is implemented with a conventional ECL master-slave DFF. Measurement result shows that the comparator has 5.8-bit of resolution with the input bandwidth of 2 GHz. Power dissipation of the core comparator is 70 mW.

Chapter 4
Current Steering DAC Architecture

_____

4. 1. Introduction

In the last few decades the communication bandwidth has evolved with an enormous speed and the requirement of high-speed data converters is directly dictated by that. In RF systems, the analog-digital interface is pushed towards the antenna, as the complex signal processing can be handled more efficiently in the digital domain.

The direct digital synthesis (DDS) technique becomes more and more popular in the mobile communication arena due to the simple control procedure rather than an analog domain phase locked loop (PLL) based signal synthesis [42][43]. The front end D/A converter (DAC) is a critical component in those systems. In high speed data links e.g. optical, radar or satellite communication systems, medium resolution (4-8 Bits) DAC with sampling rate of up to 20 GHz are going to be used [44]. Another upcoming application of high-speed medium resolution DAC can be found in ultra wideband (UWB) communication systems. Different kinds of pulse forms are used, e.g. Gaussian and its derivatives. A DAC based direct waveform synthesis (DWS) is presented in [45]. The key requirement for this application is medium resolution with sampling rate more than ~16GHz and low power.

Such ultra high-speed DACs find a new application in the highly efficient class-S power amplifier. Currently in state of the art implementations a continuous time single bit delta-sigma modulator is being used [46]. This core modulator can be replaced with a multi-bit version to achieve lower oversampling ratio and thus higher signal frequency.

For these high frequency applications current steering DAC architecture comes as an obvious choice [47] [48]. The main advantage with this architecture lies in its simplicity and the high conversion rate is achieved by employing the maximum possible parallel processing.

In this chapter brief introduction to this architecture is being presented. The static and dynamic performances of a DAC are defined with the same sets of parameters like the ADC described in chapter 2. The static accuracy is defined by the parameters integral nonlinearity

(INL) and differential nonlinearity (DNL) whereas the dynamic performances are defined with signal to noise ratio (SNR), total harmonic distortion (THD) and spurious free dynamic range (SFDR).

This chapter is organized as follows: In section 4.2 different kinds of current steering DAC architectures are presented. Various static and dynamic error sources associated with the current steering DAC have been presented in section 4.3. The most commonly used procedures to overcome the error sources in current steering DAC have been presented in section 4.4. Finally the conclusions are drawn in section 4.5.

4. 2. Current Steering DAC Architecture

The operation principle of current steering DAC can be explained as follows: There are number of current sources and switches. Depending upon the input code word X currents from the corresponding sources are directed by the switches to the output. A simple resistor or an Opamp based current to voltage converter is used to convert the output current into voltage. In this kind of architecture the static and dynamic accuracy of the DAC directly depends on the matching accuracy among the current cells.

There are number of ways to realize the current sources. According to the way of implementation of the current sources the current steering DAC can be broadly divided in to two categories; binary weighted and unary weighted, where the latter is also called directly or thermometer coded. The combination of binary and unary weighted sub-DACs; is commonly known as segmented current steering DAC. In the next sub-sections these three kind of current steering architectures are described.

4. 2. 1. Binary Weighted Current Steering DAC

In Fig. 4.1 the conceptual block diagram of an N-bit binary weighted current steering DAC is presented. These N-bits are directly used to control the N-current sources. The current sources are binary weighted i.e. the current source controlled by P^{th} input bit has a current weight of $(2^P-1)I_{unit}$, where I_{unit} is the LSB current. The main advantage of this architecture is its simplicity. It has low power dissipation and it does not require any decoding logic.

There are several major drawbacks are associated with this architecture. All of the current sources are to be matched properly otherwise static and dynamic errors occur at the output. The most critical matching requirement is associated with MSB current source, where this current source is to be matched to the sum of the rest of the current sources within $0.5\,I_{unit}$ to maintain the monotonic transfer characteristics of the DAC. The matching requirement generally dictates the upper limit of resolution.
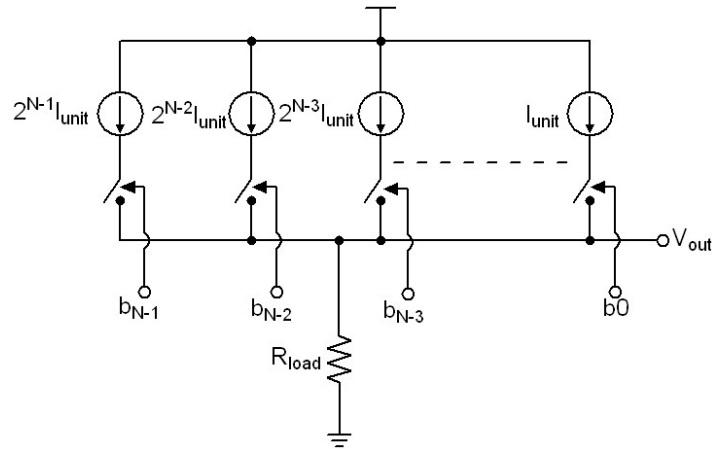
Fig. 4.1. Block diagram of binary weighted DAC

In addition to the stringent matching requirements this architecture inherently shows some dynamic errors. Among those nonidealities the most critical error is the high glitches at the output due to the current switches. As the current sources are binary weighted the height of these glitches are not constant but proportional to the weight of the current source and as well as the input bit pattern. This uncorrelated nature of the out glitches results in spurs at the output spectrum of the DAC and eventually reduces the accuracy. The worst output glitch occurs at the midcode transition i.e. in the initial state all of the current sources other than the MSB current source are connected with the output and in the next state MSB current source is connected with the output and other current sources are switched off. At this kind of transition all of the current switches are active (either switched on or off), which results in the high glitch at the output.

4. 2. 2. Unary weighted Current steering DAC

In Fig. 4.2 a simplified block diagram of an N-bit unary weighted current steering DAC has been presented. Unlike the binary weighted DAC, all the current sources have the same weight (I_{unit}). The input N-bit binary code is converted into thermometer code by a

thermometer decoder, which generates 2^N-1 number of control signals. When the input digital input increases by 1LSB, one more current source is switched to the output. Thus the analog output is always increasing as the digital input increases. Hence, the monotonicity is always guaranteed in this architecture.

In addition, there are several other advantages of unary weighted DAC compared to its binary weighted DAC. The matching requirement is much relaxed in this architecture. 50% matching of unit current source is good enough for DNL ≤ 0.5LSB. But unfortunately INL error can be higher. Several techniques are used to reduce the INL error. Some of the state of the art techniques are explained in the later part of this chapter (Section 4.4). At midcode transition (as explained in section 4.2.1) only one additional unit current source is switched to the output. Thus the midcode transition glitch is greatly reduced.
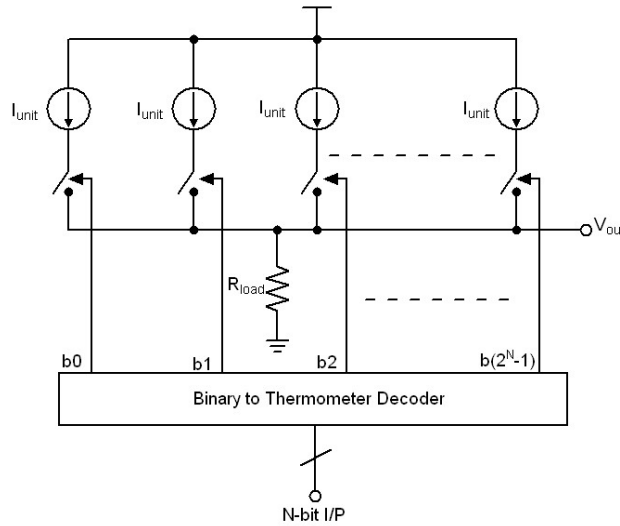
Fig. 4.2. Block diagram of binary weighted DAC

One of the main advantages of unary weighted architecture is that, the output glitches hardly contribute to the nonlinearity. This is because the magnitude of the output glitch depends upon the number of current sources switch to the output. Since the number of current sources switched to the output is proportional to the amplitude of the input signal steps, hence it does not increase the nonlinearity [49].

This architecture comes with a higher area overhead, due to large number of current cells and the combinational logic. The number of unit current cells increases exponentially with the increasing input bits and so does the complexity of thermometer decoder, which generally imposes the upper limit of this architecture.

4. 2. 3. Segmented Current Steering DAC

A fully unary weighted DAC guarantees the monotonicity and minimal glitches. However, for high resolution this is not feasible, as this architecture takes large die area. In binary weighted DAC, the area requirement is relaxed but it has higher nonlinearity. To get the best from both of the architectures, most of the high speed high resolution current steering DACs are implemented with segmented current steering architecture. A simplified block diagram of such N-bit segmented current steering DAC is presented in Fig. 4.3. This N-bit DAC is divided into two sub-DACs. M-bit LSBs are implemented with the binary weighted architecture, whereas rests of (N-M)-bit are realized with unary weighted DAC architecture. The input binary bits of the unary weighted DAC are converted into thermometer code by a binary to thermometer decoder, which has significant delay. On the other hand the binary weighted DAC does not require any decoding logic. To equalize the delay of unary and binary weighted sub-DAC outputs, the inputs of the binary weighted sub-DAC are delayed by the delay equalizer block.
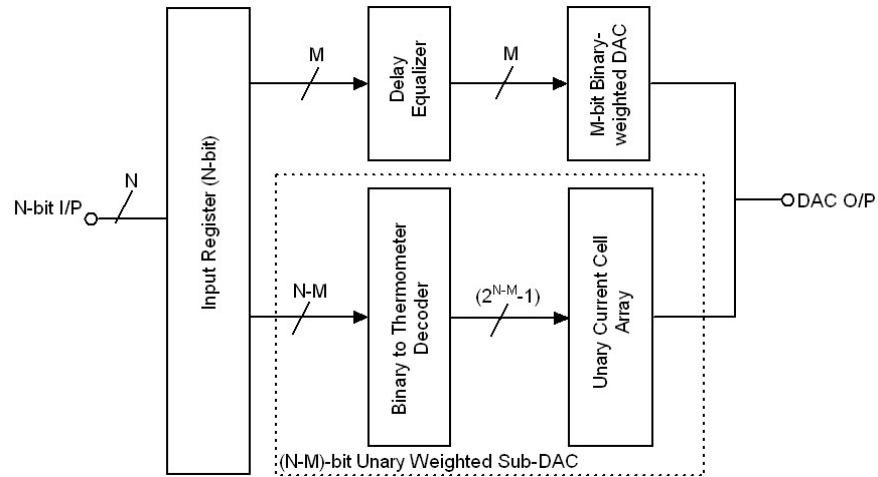
Fig. 4. 3. Simplified block diagram of segmented current steering DAC

In [49] a mathematical analysis of the percentage of segmentation with the area and linearity of the DAC has been presented, where the full binary weighed implementation is referred as 0% segmentation and full unary weighted implementation is referred as 100% segmentation. With the increasing percentage of segmentation the area of the DAC is increased exponentially and the static accuracy of the DAC get improved. The DNL error reduces almost linearly with the increasing percentage of segmentation. Afterward to fulfill

the INL requirement the area is further increased. Finally the chip area is dominated by the size of the thermometer decoder.

4. 2. 4. R-2R ladder DAC

(a)

(b)
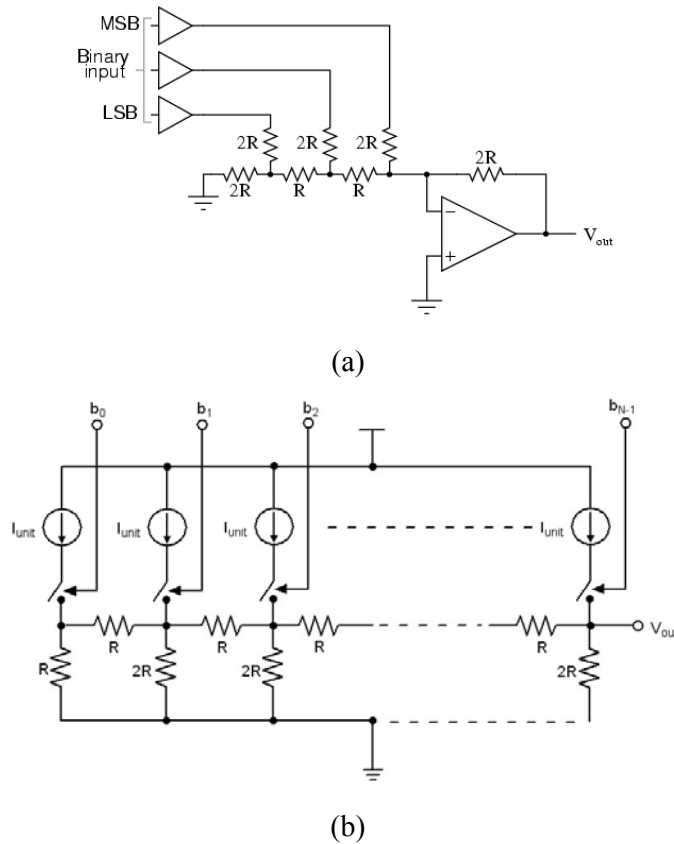
Fig. 4. 4. (a) conventional (b) improved R-2R ladder DAC architecture

A basic R-2R resistor ladder network is shown in Figure 4.4a. The digital inputs or bits range from the most significant bit (MSB) to the least significant bit (LSB). The bits are switched between either 0V or V_{REF} and depending on the state and location of the bits the output voltage V_{OUT} will vary between 0V and V_{REF} minus one LSB's voltage. The main problem of this architecture arises in the switches. Depending upon the position of resistor the current trough the switch varies as well as the switching time. In high-speed application this variation in the switching time results in harmonic distortion. A improved version of R-2R DAC is presented in Fig. 4.4b. This R-2R ladder DAC can be considered as a special kind of current steering architecture where all of the current cells are having same weight and the binary weighing operation is implemented with the resistive ladder. This architecture is suitable for processes, which are capable of implementing highly linear resistors. The R-2R

ladder architecture is shown in Fig. 4.4b. All current sources have the same weight (I_{unit}) and the switches are controlled by the N-bit digital input (Similar to the binary weighted architecture). Since all of the slices consisting of a current source, a switch, a resistor R, and 2R resistor, a modular layout can be done. This enhances the matching among different components. As the current sources are all equally large, a special current source trimming techniques can be applied.

Looking from the output (from the left to the right in the figure) the impedance is always R. The current switched by the LSB (b_0) takes the longest time delay to appear at the output whereas the current of the MSB source appears at the output with the shortest delay. The time-delay between the switches of the MSB to LSB generates glitches for this architecture [4]. In the R-2R ladder architecture shown in Fig. 4. 4, there is the same amount of current through all switches, which makes the design of the switches simpler, and current switching dynamics similar. However, the internal voltage nodes are varying with time and therefore the current sources will have varying terminal voltages, hence resulting in nonlinearity and distortion.

4. 3. Error sources in Current steering DAC

Depending upon the architecture, current steering DACs are composed of number of binary or unary weighted current cells, which include the current source and the differential current switch. Any nonideality that occurs in these current cells directly influences the static and dynamic characteristics of the full DAC. The matching accuracy among the current sources has direct impact on the static accuracy (INL and DNL) of the DAC whereas the instantaneous output impedance and the switching delay deteriorate the dynamic performance (SFDR, THD). In this section an attempt has been made to explain the main static and dynamic error sources.

4. 3. 1. Static Error Source

As mentioned earlier the static accuracy of the current steering DAC directly depends on matching among the current sources. The error caused by the process variation (area, threshold voltage, oxide thickness) can be broadly divided into two categories; the random and graded variations [50]. The random process parameter variation can be considered to be a

statistical process and generally it has Gaussian probability distribution. On the other hand the graded variations are systematic errors (linear, quadratic or higher order).

Due to the mismatch of the current source transistors the INL specification of different DACs produced in the same process technology varies randomly. To predict the INL specification within a certain boundary a well accepted parameter called INL_{yield} is introduced. This particular figure of merit is defined as the percentage of functional DAC with an INL specification smaller than 0.5LSB.

The first analytical formulation to determine the INL_{yield} was proposed in [51].

$$INL_{yield} = \prod_{i=2}^{2^N-1} erf\left(\frac{Q_i}{\sqrt{2}}\right) \qquad (4.1)$$

where,

$$\frac{1}{Q_i} = 2^{N+1}\left[\frac{\overline{Z_X}\left(1-\overline{Z_X}\right)}{2^N-1}\right]^{\frac{1}{2}}\left[\frac{\sigma(I_{LSB})}{I_{LSB}}\right] \qquad (4.2)$$

N is the number of input bits. $\overline{Z_X}$ is the normalized mean output for the N-bit input code X and $\sigma(I_{LSB})$ is the is the standard deviation of the unit current I_{LSB}.

In equation 4.1 it is assumed that all of the unit current cells are uncorrelated and the total error probability for the current cells can be found by multiplying the errors of the individual current cells. But in reality these current cell errors are not uncorrelated. Thus equation 4.1 leads to worst case situation. An improvement of equation 4.1 is proposed in [52]. Here the mid-code transition is viewed as the most critical event since in binary weighted DAC implementation this transition has the largest probability of generating output error (see section 4.2.1). The modified INL_{yield} can be expressed as,

$$INL_{yield} = \prod_{i=2^{N-1}-1}^{2^{N-1}} erf\left(\frac{Q_i}{\sqrt{2}}\right) \qquad (4.3)$$

where Q is defined according to equation 4.2. The equation 4.3 gives an optimistic value for the INL_{yield}, as it considers the only a single transition but there is a probability that the error occurs in the other transitions. However from equation 4.1 and equation 4.3 the upper and lower limit of the area of a unit current source can be predicted. For a given INL_{yield} value the $\sigma(I_{LSB})$ can be calculated, which can be used to calculate the area of the current source. A well accepted relation is proposed in [53] to relate the area and the matching error of the

MOS current sources, which include almost all sources of random and graded variations. It, can be expressed as,

$$\frac{\sigma^2(I_{LSB})}{I_{LSB}} = \frac{4\sigma^2(V_{T0})}{(V_{GS} - V_{T0})^2} + \frac{A_\beta^2}{W \cdot L} \tag{4.4}$$

where, W and L are the width and length of the MOS transistor. V_{T0} is the threshold voltage of the MOS transistor and the V_{GS} is gate to source voltage. A_β is a constant for any particular process technology, which relates to the different matching coefficients e.g. the oxide thickness, mobility, length and width variation of a MOS transistor. Equation 4.4 implies that the matching performance of the MOS current sources can be improved by increasing the area of it and as well as the drain to source overdrive voltage.

By using equations 4.1, 4.3 and 4.4 the area of a unit current source can be calculated which would provide the best and worst case INL_{yield} performance. But to calculate the optimum INL_{yield} the most commonly used process is Monte Carlo approach [54], [55].

4. 3. 2. Dynamic Error Sources

In current steering DAC any mismatch among the current sources leads to static errors in the output. But there are some other effects generally cause time and input code dependent nonlinearity, which in turn deteriorate the dynamic performances (THD, SFDR, SNDR) of a DAC. In the following sub-sections the main sources of dynamic errors have been explained.

4. 3. 2. 1. Finite Output Impedance

In Fig. 4.1 a simplified schematic of N-bit binary weighted current steering DAC is presented. The binary weighted current sources are implemented with parallel combination of unit current source I_{unit} and the current switches are directly controlled by N-bit input word. At a certain time point t=nT, the output current is represented as $I_{out}(nT)$, where T is the sampling time period. The N-bit input digital word is denoted by, $X(nT)=\{b_{N-1}, b_{N-2},\ldots,b_1,b_0\}$, where $b_0,\ldots\ldots,b_N$ are input bits. Thus $I_{out}(nT)$ can be expressed as,

$$I_{out}(nT) = I_{unit}b_0(nT) + 2I_{unit}b_1(nT) + \cdots\cdots\cdots + 2^{N-1}I_{unit}b_{N-1}(nT) \tag{4.5}$$

whereas, the input word X(nT) (for simplicity X(nT) will be represented as X) can be represented as,

$$X(nT) = X = 2^{N-1}b_{N-1}(nT) + 2^{N-2}b_{N-2}(nT) + \cdots\cdots\cdots + 2b_1(nT) + b_0(nT) \quad (4.6)$$

Combining Equations 4.5 and 4.6, $I_{out}(X)$ can be expressed as,

$$I_{out}(X) = I_{unit}X \quad (4.7)$$

In Fig. 4.5 the small signal equivalent circuit of a current source along with the load resistance (R_{load}) is presented. For simplicity the current switch is assumed to be ideal. The finite output impedance of the current source and the parasitics associated with interconnects have great influence on the dynamic performance of current steering DAC. As shown in Fig. 4.5 the nonideal current source can be modeled as a parallel combination of ideal current source (I_{out}) and the finite output resistance ($1/G_{out}$). In a particular time instant when only one of the current cells is connected to the output load, then the load current (I_{load}) can be expressed as,

$$I_{load} = \frac{I_{out}}{1 + R_{load}G_{out}} + \frac{G_{out}V_{DD}}{1 + R_{load}G_{out}} \quad (4.8)$$

As the equation 4.8 implies, the effect of finite output conductance (G_{out}) of the current source introduces gain error and as well as offset error. But it does not affect the linearity of the DAC when the output impedance of the current source remains constant.
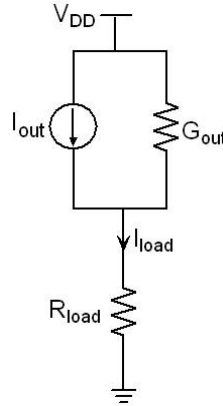
Fig. 4.5. Small signal equivalent model of unit current source

But in reality the output impedance of current steering DAC depends on the input word X. Assuming the input dependent output conductance is denoted by $G_{out}(X)$, the input dependent load current $I_{load}(X)$ can be expressed as follows,

$$I_{load}(X) = \frac{I_{out}(X) + V_{DD}G_{out}(X)}{1 + R_{load}G_{out}(X)} \quad (4.9)$$

73

The input dependent output conductance ($G_{out}(X)$) is the parallel combination of the number of unit current cells switched to the load (R_{load}), which is directly controlled by the input word X. If a unit current cell has output conductance of G_{unit} then,

$$G_{out}(X) = 2^{N-1}b_{N-1}(nT)G_{unit} + 2^{N-2}b_{N-2}(nT)G_{unit} + \cdots\cdots + 2b_1(nT)G_{unit} + G_{unit}$$
$$= G_{unit}X \qquad (4.10)$$

If the ratio of the load resistance (R_{load}) and the output resistance of the unit current source ($1/G_{unit}$) is defined as ρ,

$$\rho = R_{load} \bullet G_{unit} = \frac{R_{load}}{R_{unit}} \qquad (4.11)$$

where R_{unit} is the output resistance of the unit current source. Now combining equations 4.7, 4.9 and 4.11 input dependent load current can be rewritten as,

$$I_{load}(X) = \frac{I_{unit} + G_{unit}V_{DD}}{\rho}\left(1 - \frac{1}{1+\rho X}\right) \qquad (4.12)$$

The input signal X is assumed as a single tone sinusoidal,

$$X = V_{DC} + V_a \sin\alpha + Q \qquad (4.13)$$

where, V_{DC} is the dc level of the input sinusoidal, V_a is the amplitude of input sinusoidal and α is the normalized input frequency. Q is the quantization noise, which can be assumed as white noise for large number of input bits.

The SFDR of a data converter is defined as the difference between the fundamental and largest uncorrelated frequency component within the output band of interest of a data converter. By replacing the input signal (X) within parenthesis in equation 4.12 with equation 4.13 and expanding as converging Taylor series an estimation of this SFDR is presented in [56], and it is reproduced in the following equation,

$$SFDR = \left[\frac{V_{DC} + R_{ratio}}{V_a} + \sqrt{\left(\frac{V_{DC} + R_{ratio}}{V_a}\right)^2 - 1}\right]^2 \qquad (4.14)$$

Where, $R_{ratio} = 1/\rho = R_{unit}/R_{load}$.

In most of the applications the input dc level (V_{DC}) and the input sinusoidal amplitude (V_a) are equal. In that case equation 4.14 can be rewritten as,

$$SFDR = \left(1 + \frac{R_{ratio}}{V_a}\left[1 + \sqrt{1 + \frac{2V_a}{R_{ratio}}}\right]\right)^2 \tag{4.15}$$

As equation 4.15 implies, the SFDR performance can be improved by two means: by increasing the input amplitude or by increasing the output resistance of unit current source (R_{out}) for a given load resistance. Generally the input amplitude is fixed for a given application so the most attractive way to improve the SFDR is to enhance the R_{unit}.

4. 3. 2. 2. Asynchronous Switching

The delay related nonlinearity is one of the main contributors to the bad dynamic behavior of a high speed high resolution DAC. In Fig. 4.6 a most commonly used floorplan of a unary weighted DAC is presented. The input signal is converted into thermometer code and the outputs of the thermometer decoder directly control the unit current cells. All of the unit current cells are placed in a matrix format. A simplified schematic of such a unit current cell is presented in Fig. 4.7. It is composed of two parts, a current switch with a unit current source and the latch, which controls the current switching.
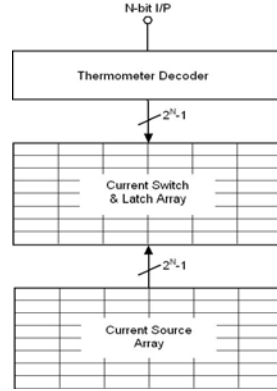
Fig. 4.6. Commonly used floorplan for unary weighted DAC

In unary weighted current steering DAC architecture the number of current cells will increase exponentially with the increasing resolution. All current cells are controlled by the input latch (as shown in Fig. 4.7). These latches in turn are synchronized by a global clock input, which is connected to a clock input pad. When the accuracy of the DAC increases, it becomes more and more difficult to let all these current sources have the same delay from the clock pad keeping a reasonable chip size of the DAC. The delay from the clock pad to the latch as well as the delay from the individual outputs of the current cells to the output pads do not depend on the output values rather the position of the current cells in the matrix. These

kinds of delays are termed as cell dependent delays. Considering the current cell shown in Fig. 4.7 the differential output signal does not reduce the cell dependent delays as both of the differential outputs are shifted by the same amount. This delay results in the higher second order harmonic at the output of the DAC.
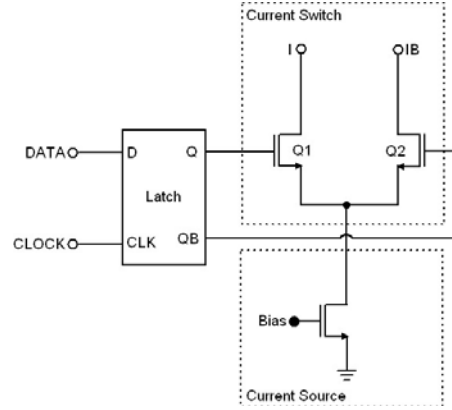
Fig. 4.7. Simplified schematic of unit current cell

If the ith current cell has delay of d_i. At any point of time t=nT, the output current of ith current cell ($I_{i,n}(t-d_i)$) is given by,

$$I_{i,n}(t-d_i) = 1 - exp\left(-\frac{t-nT-d_i}{\tau}\right) \quad where, nT \le t \le (n+1)T \tag{4.16}$$

where, T is the sampling time period and τ is the time constant decided by the output load of the DAC. For simplicity the amplitude of unit current cell is assumed to be 1. Thus the distortion caused by delay of ith current cell ($\delta_{i,n}$) is given by,

$$\delta_{i,n}(t) = I_{i,n}(t) - I_{i,n}(t-d_i) \approx \frac{d_i}{\tau} exp\left(-\frac{t-nT}{\tau}\right) \quad where, \quad nT \le t \le (n+1)T \tag{4.17}$$

Assuming the DAC input is a sinusoidal signal,

$$f_{in}(t) = 2^{N-1}\left[sin(\omega_0 t)+1\right] \tag{4.18}$$

where, N is the number of input bits of the DAC. In addition it is assumed that the input sinusoidal has an amplitude, which is half of the output full scale voltage of the DAC with a DC level at the middle of the output full scale voltage. Under the consideration that the resolution of the DAC is high, the quantization noise can be ignored. Thus the ideal output of the DAC in a period of t=nT to t=(n+1)T can be expressed as,

$$A_n \approx 2^{N-1}\left[sin(\omega_0 nT)+1\right] \tag{4.19}$$

$$A_{n+1} \approx 2^{N-1}\left[sin(\omega_0(n+1)T)+1\right] \tag{4.20}$$

76

During this period the total DAC distortion is,

$$\Delta_n(t) \approx \sum_{i=A_n}^{A_{n+1}} \delta_{i,n}(t)$$

$$\approx \sum_{i=A_n}^{A_{n+1}} \frac{d_i}{\tau} \exp\left(-\frac{t-nT}{\tau}\right) \bullet G_T\left[t - \left(\frac{2n+1}{2}\right)T\right] \tag{4.21}$$

where, $G_T(t)$ is the square function used to calculate the DAC distortion in the given time window and is defined as follows,

$$G_T(t) = \begin{cases} 1 & where, \quad -T/2 \le t \le T/2 \\ 0 & else \end{cases} \tag{4.22}$$

For further calculation the delay (d_i) is considered to be linearly distributed i.e.

$$d_i = a \bullet i \tag{4.23}$$

where a is a constant which corresponds to the unit delay associated with single current cell. Thus depending upon the switching sequence the delay increases linearly. In addition the current cells are switched on in the same order as their delay increases. Thus to calculate the distortion for the time window t=nT to t=(n+1)T, equation 4.23 can be replaced in equation 4.21 and after calculation can be rewritten as [57],

$$\Delta_n(t) = \frac{a}{\tau} \exp\left(-\frac{t-nT}{\tau}\right) \bullet G_T\left[t - \left(n + \frac{1}{2}\right)T\right] \bullet \frac{1}{2}\left[A_{n+1}^2 - A_n^2 + A_{n+1} + A_n\right] \tag{4.24}$$

Defining the function M(t) as,

$$M(t) = \frac{a}{\tau} \exp\left(-\frac{t}{\tau}\right) \bullet G_T\left[t - \frac{T}{2}\right] \tag{4.25}$$

The equation 4.24 can be rewritten as,

$$\Delta_n(t) = M(t-nT) \bullet \frac{1}{2}\left[A_{n+1}^2 - A_n^2 + A_{n+1} + A_n\right] \tag{4.26}$$

In equation 4.26, M(t-nT) appears as the amplitude component for different frequencies. Replacing the values of A_n and A_{n+1} with the equations 4.19 and 4.20 respectively and rearranging, the distortion in nth time period can be expressed as,

$$\Delta_n(t) = \frac{M(t-nT)}{2}\{2^{2N-2}\sin(2\omega_0 nT + \omega_0 T)\sin\omega_0 T + \left(2^{2N-1} + 2^{N-1}\right)\sin(\omega_0(n+1)T) \\ + \left(2^{N-1} - 2^{2N-1}\right)\sin\omega_0 nT + 2^N\} \tag{4.27}$$

In equation 4.27 the arguments of sin functions are in discrete domain. Those can be changed with continuous time representation, i.e. $\sin(\omega_0 nT)$ with $\sin(\omega_0 t)$. Then the multiplication with

77

the δ(t-nT) makes is discrete once again. Then the overall distortion of the DAC can be expressed as,

$$
\begin{aligned}
\Delta(t) \quad &= \sum_{n=-\infty}^{\infty}\Delta_n(t) \\
&= \frac{M(t)}{2} * \left\{ \left[2^{2N-2} \sin(2\omega_0 t + \omega_0 T)\sin\omega_0 T + \left(2^{2N-1} + 2^{N-1}\right)\sin(\omega_0 t + \omega_0 T) + 2^N \right. \right. \\
&\quad \left. \left. + \left(2^{N-1} - 2^{2N-1}\right)\sin\omega_0 T \right] \sum_{n=-\infty}^{\infty}\delta(t-nT) \right\}
\end{aligned}
\tag{4.28}
$$

where, $*$ is the convolution operator.

For high resolution DAC the distortion components are relatively small compared to the fundamental output component. Therefore the distortion is mainly contributed by the second order harmonic component and the distortion due to the second order harmonic can be approximated as,

$$
\Delta_{II}(t) = \frac{M(t)}{2} * \left[2^{2N-2} \sin(2\omega_0 t + \omega_0 T)\sin\omega_0 T \sum_{n=-\infty}^{\infty}\delta(t-nT) \right]
\tag{4.29}
$$

Applying Fourier transform to equation 4.29 and simplifying the second order harmonic component can be expressed as [57],

$$
\left|\Delta_{II}(2\omega_0)\right| = \frac{2^{N-4}\, d_{max}\,\omega_s\, \sin\!\left(\dfrac{2\pi\omega_0}{\omega_s}\right)}{\sqrt{1 + (2\omega_0\tau)^2}}
\tag{4.30}
$$

Where, $\left|\Delta_{II}(2\omega_0)\right|$ is the amplitude of the second order harmonic component. d_{max} is the maximum delay of the current cells i.e.

$$
d_{max} = a \bullet 2^N
\tag{4.31}
$$

ω_s corresponds to the angular frequency of sampling time period (T).

From equation 4.30 it is evident, that the second order harmonic increases with the sampling frequency and the maximum delay difference among the current cells. This harmonic component can be reduced by increasing the time constant at the output (τ), which appears as a low pass filter at the output. But this time constant cannot be increased after a certain value otherwise the output bandwidth of the DAC is reduced. For a given sampling frequency the most efficient way to reduce the distortion by minimizing the maximum delay (d_{max}) with efficient floorplanning of the current cells.

4. 3. 2. 3. Current Switch Non-idealities

(a)

(b)
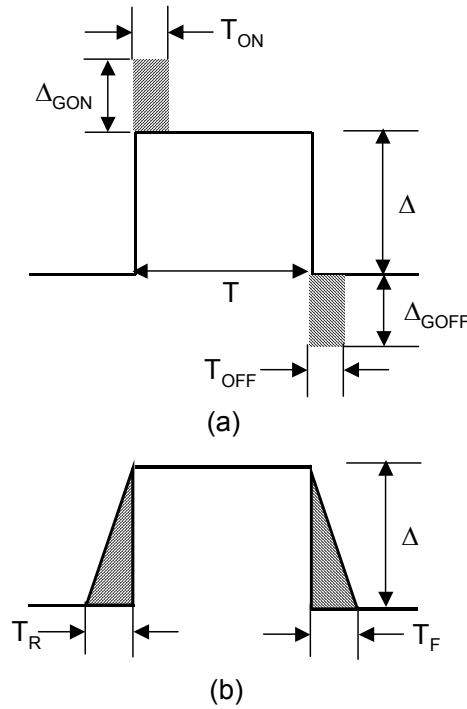
Fig. 4.8 (a) Representation of output glitch due to the charge injection and clock feedthrough of current switch

(b) Finite rise and fall time for the built-in-time constant of the current switch

In the previous sub-section the error due to switching delay and its effects has been explained. In this sub-section the nonidealities associated with the current switches are presented. The main error sources of any unit current switch can be identified as the charge injection, clock feedthrough and the built-in-time constant. In Fig 4.8a the effects of charge injection and clock feedthrough of an individual current switch are presented whereas the Fig. 4.8b represents finite rise and fall time of the current switch due to the built-in-time constant of the current switch.

The charge injection and the clock feedthrough of the current switches cause output glitches and these output glitches can be approximated as two rectangular pulses (shaded region of Fig. 4.8a). The widths of these glitches are T_{ON} and T_{OFF} during the current switch turns on and off respectively. The height of these turn on and turn off pulses are given by ΔG_{ON} and ΔG_{OFF}. The ideal output pulse has a time period of T with height of Δ. Thus the normalized area of the turn on (ε_{ON}) and turn off (ε_{OFF}) pulses can be expressed as,

$$\varepsilon_{ON} = \frac{\Delta G_{ON} T_{GON}}{\Delta \cdot T} \tag{4.32}$$

$$\varepsilon_{OFF} = \frac{\Delta G_{OFF} T_{OFF}}{\Delta \cdot T} \tag{4.33}$$

To calculate the error due to the turn on and turn off glitches a unary weighted DAC with input sinusoidal signal as presented in equation 4.18 is assumed. Further more the radian frequency of the input signal is approximated as unity. If a high oversampling ratio (OSR) is considered then the output of the DAC can be approximated as a triangular signal. For one of the half cycle ($0 \leq t < \pi$) of the input, at any point of time t the error would be negative and for other half cycle the switching error would be positive. Under this circumstance the error (e(t)) due to the output glitches for the full cycle of the input signal is given by [58],

$$e(t) = \begin{cases} 2^{N-1} \varepsilon_{OFF} \left[\sin(t) - \sin(t - T) \right] & for, \ 0 \leq t < \pi \\ 2^{N-1} \varepsilon_{ON} \left[\sin(t) - \sin(t - T) \right] & for, \ \pi \leq t < 2\pi \end{cases} \tag{4.34}$$

Performing Fourier transform over equation 4.34 with high OSR, the even order components can be presented as [18],

$$\Delta_{2n} \approx 20 \log_{10} \left(\frac{2^{N-1} \sin\left(\dfrac{\pi}{OSR}\right) \bullet \left| \varepsilon_{OFF} - \varepsilon_{ON} \right|}{\pi(2n+1)(2n-1)} \right) \tag{4.35}$$

Where Δ_{2n} is the even order harmonic components and n=1, 2, 3,……

From equation 4.35 it can be observed that any asymmetry in the turn on and turn off glitches in the current switches results in higher even order harmonic components and it is proportional to difference of the areas of these glitches.

4. 4. Techniques to Enhance the Accuracy of Current Steering DAC

To enhance the static and dynamic performance of the current steering DAC several techniques are used. Those techniques can be divided into three broad categories. Different layout techniques are used to reduce the random mismatches and the graded errors. Dynamic element matching can be used to randomize the INL error due to the mismatch among the current cells. Special current cell calibration techniques can be used to enhance the accuracy matching accuracy of different current cells. In the following section these techniques are briefly described.

4. 4. 1. Layout Technique

In section 4.3.1 it has been explained that the random and graded process parameter variation causes severe nonlinerity in the static output characteristics of the current steering DAC. To cope up with the symmetrical and graded errors caused by temperature, process parameter and electrical gradients, generally common centroid layout technique is used.

One of such an example can be found in [59]. A 10-bit DAC has been implemented with a segmented current steering architecture. The 5LSBs are implemented with binary weighted architecture and rests of the 5MSBs are realized with unary architecture. The unary weighted sub-DAC has 31 unit current sources. All of the unary current sources are divided into four units and these units are placed in four different quadrants. This technique is generally known as double centroid layout technique. To place the 31units into each quadrant, a 6X6 matrix format is used. In addition to that four rows and four columns of dummy cells are used around the core area to reduce the edge effect. A pictorial representation of such a double centroid layout technique is shown in Fig. 4.9. In this implementation a 4-bit unary DAC implementation is presented. The shaded cells are dummy units, which are used to reduce the edge effect.

		15	14									14	15		
			12	10					10	12					
		8	4	2	6			6	2	4	8				
		5	1	3	7			7	3	1	5				
			9	11					11	9					
					13	13									
					13	13									
			9	11					11	9					
		5	1	3	7			7	3	1	5				
		8	4	2	6			6	2	4	8				
			12	10					10	12					
	15	14									14	15			

Fig. 4.9. Floorplan of double centroid unary current source array

An extension of double centroid layout technique can be found in [60]. Here also the current source array is divided into four quadrants and each of these quadrants is further divided into four sub-quadrants. Thus instead of four components every unit current source is composed of sixteen components. By using this layout technique an intrinsic accuracy of 14-bits has been achieved in [60].

An improved layout scheme is proposed in [61] to cope up with the gradient error for the unary DACs. For N-bit unary weighted DAC implementation, 2^N-1 number of current sources are divided into 2^N-1 number of equal components. Therefore the current sources form a $(2^N-1)\text{X}(2^N-1)$ matrix. A single current source has exactly one component in every row. As an example 8-unit current cells are implemented and the floorplan of this current source matrix is presented in Fig. 4.10. This current source matrix has 8 rows and 8 columns and the current source components are represented by $\{0,1,2,\ldots,7\}$. Elements with the same number belong to the same current source. This layout technique is very useful to reduce the linear gradient error along the x and y-axes. As there is one component of every current source in each row. As a result of it the summation of the error along the x-axis is same for all the current sources. The same argument holds for any linear gradient error along the y-axis. But one of the drawbacks of this layout scheme is that it cannot reduce the quadratic gradient errors.

1	7	5	3	0	6	4	2
7	0	2	4	6	1	3	5
5	3	1	6	4	2	0	7
2	4	6	0	3	5	7	1
0	6	4	2	1	7	5	3
6	1	3	5	7	0	2	4
4	2	0	7	5	3	1	6
3	5	7	1	2	4	6	0

Fig. 4.10. Linear gradient error reducing layout scheme

An improved floorplan for the 8 current sources to reduce the quadratic gradient errors is shown is Fig. 4.11. The size of the matrix is doubled here compared to the layout procedure shown in Fig. 4.10 with the mirrored copy of it. In [61] it is shown, this layout scheme can practically cancel the quadratic gradient errors. It only produces an offset, which does not contribute to the linearity of the DAC.

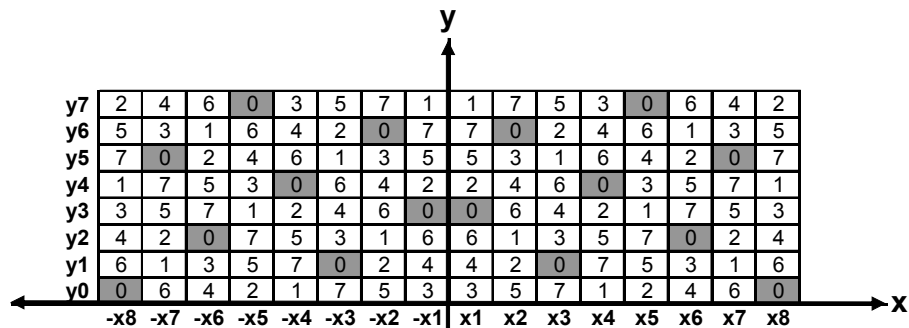| y																
y7	2	4	6	0	3	5	7	1	1	7	5	3	0	6	4	2
y6	5	3	1	6	4	2	0	7	7	0	2	4	6	1	3	5
y5	7	0	2	4	6	1	3	5	5	3	1	6	4	2	0	7
y4	1	7	5	3	0	6	4	2	2	4	6	0	3	5	7	1
y3	3	5	7	1	2	4	6	0	0	6	4	2	1	7	5	3
y2	4	2	0	7	5	3	1	6	6	1	3	5	7	0	2	4
y1	6	1	3	5	7	0	2	4	4	2	0	7	5	3	1	6
y0	0	6	4	2	1	7	5	3	3	5	7	1	2	4	6	0
	-x8	-x7	-x6	-x5	-x4	-x3	-x2	-x1	x1	x2	x3	x4	x5	x6	x7	x8

Fig. 4.11. An improved linear gradient error reducing layout scheme

82

The layout techniques mentioned in this section are very useful to increase the linearity of the DAC. But it comes with higher area overhead and as well as complex routings are required which make this technique unattractive for high speed applications.

4. 5. 2. Dynamic Element Matching

The principle of dynamic element matching can defined as a process to enhance the matching accuracy of different poorly matched devices by performing the time average operation over these components. This technique can be used in current steering DAC architecture to increase the linearity.
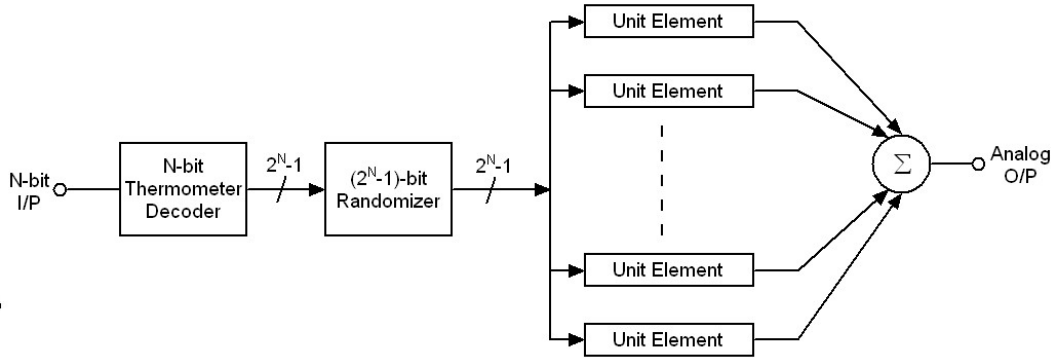
Fig. 4.12. Architecture of dynamic element matching unary weighed DAC

This dynamic element matching DAC topology can be constructed using any unary weighted N-bit D/A converter in which the Pth output level is generated by activating P number of approximately equal-valued elements, typically resistors, capacitors or current sources and summing up their charge or current or voltage (see Fig. 4.12). Dynamic element matching is implemented by choosing different elements to represent the Pth level as a function of time. The "randomizer" block decides, which elements will be used to represent the Pth level on each clock cycle. The goal of this approach is to convert the error due to element mismatch from a dc offset into an ac signal of equivalent power, which in an oversampling converter, can be partially removed by filtering. Even when the input is constant, the error is a wide-band noise signal. With ideal randomization, a mismatch between the unit elements would be converted into a white-noise signal with zero mean error and a variance equal to the root-mean-square (rms) error between the individual unit elements.

First, let us consider the linearity of such DAC. For a fixed input code of X, each element is active, on average, X out of every $M=(2^N-1)$ clock cycles (where 2^N-1 is the total number

of elements for N-bit unary weighted DAC). Therefore, each element of the DAC acts individually as a duty-cycle modulator and the integral linearity is limited only by the product of the fractional element mismatch error (ΔE/E) and the fractional clock jitter (ΔT/T) [1], [62]. A second practical limit on the integral linearity results because there is normally a small change in the charge (or current) transferred by each element as a function of the number of elements active. With careful choice of DAC topology and the use of a precision clock, extremely high dc integral linearity can be achieved, even when the elements match very poorly. However, the element mismatch now appears as an ac noise signal added to the DAC output.

If small scale error factors are ignored, the maximum noise signal n(t) varies in a parabolic fashion from zero at either zero or full scale to a maximum at half of full scale. At this maximum, n(t), relative to the internal DAC full scale M, is [63]

$$rms\left[\frac{n(t)}{M}\right] = \frac{rms\left[\dfrac{\Delta E}{E}\right]}{2\sqrt{M}} \tag{4.36}$$

The randomizer design can be a bottleneck in this kind of DAC implementation. The randomizer connects M outputs from the thermometer decoder to the M current switch elements. The number of possible connections could be factorial M (M!). One simple approach to randomizing over a subset of possible connections would be to have an M-port barrel shifter, which rotates one increment after each clock. This represents only M of the factorial M possible permutations. This approach would completely reduce successive output errors only if the mismatch between elements were independent of the element's position on the die. Unfortunately adjacent elements are much more likely to match than distant elements due to gradients in process parameters across the wafer.
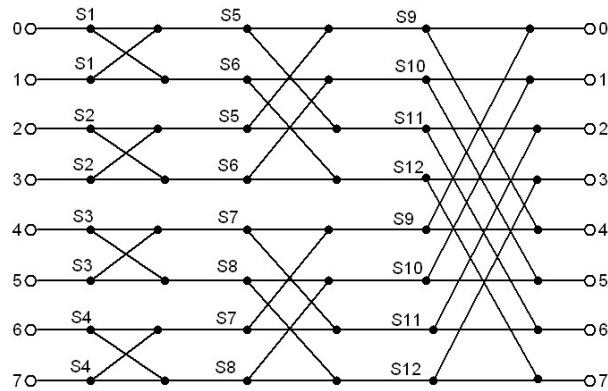

Fig. 4.13. Example of three stage butterfly randomizer

A compromise between these two extremes can be found in the "butterfly" randomizer [63]. The butterfly randomizer circuit consists of a series of butterfly networks coupling the inputs to the outputs (see Fig. 4.13). In order that any input can be connected to any output, the number of butterfly stages should be at least equal to the number of bits in the DAC. More butterfly stages can be added if it is necessary to cover a larger fraction of possible connections. A pseudorandom sequence generator would normally be used to generate the random control sequences for the butterfly switches [63].

One of the major drawbacks of the above randomizer is the high output glitch due to the large number of current switching at the output. An improved randomization technique is proposed in [64]. In this technique the randomization is done only in set of current cells, which require to change their states (i.e. from off to on state and opposite).

In spite of all these measures the dynamic element matching is suitable for low speed DACs with high oversampling ratio. Therefore technique is not very suitable for the Nyquist DACs. In addition to that the design of the randomizer becomes more power hungry in high speed applications.

4. 4. 3. Current Cell calibration technique

In the previous sections two different techniques to improve the static and dynamic accuracy of the DAC have been presented. The layout technique is very useful to reduce the effect of process parameter gradient error and randomize this error over the full chip. But it requires complex routing of the clock and output lines, which is sometimes not very suitable for high speed DAC particularly when the conversion rate is tens of gigahertz range. The clock and output path lengths come as a critical factor in those implementations. On the other hand the dynamic element matching technique does not provide a good solution in that high speed regime. Implementation of the randomizer in this frequency range is difficult and power hungry. As an alternative of the aforementioned techniques background or real-time calibration technique is widely used in the high speed DAC [67].

In this technique calibration is applied to MSB current cells of a segmented current steering DAC. The block diagram of such a calibration loop is presented in Fig.4.14. In this figure it is assumed that the current source I has to be matched to the reference current I_{ref}. A variable current source I_{cal} is added in parallel with I_{ref}. The accuracy of the matching depends on the accuracy of the current comparator and the variable current source.
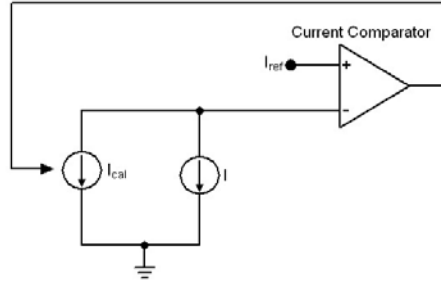
Fig. 4.14. Block diagram of current source calibration

In general the variable current source can be implemented with a precise current steering DAC. The static accuracy of the calibration DAC restricts accuracy of the full calibration loop. As an improvement, principle of calibration with non-binary weighted current sources was presented in [68]. The main advantage of this technique is that it can tolerate higher mismatch for a given accuracy. [68] concerns offset calibration of an amplifier. A B-bits binary counter is used to control a number of B non-binary weighted current cells. In the process of calibration the binary counter monotonically increases or decreases its count value to reduce the error within allowable range. In worst case the calibration process can take as long as 2^B clock cycles.

The principle of non-binary weighted calibration technique is used for the calibration of current sources. In the calibration technique, the main improvement is done in the speed of the error tracking procedure. Successive approximation technique is used to accomplish the error tracking. The block diagram of such calibration loop is presented in Fig. 4.15. The successive approximation register (SAR) controls the non-binary weighted calibration DAC through an N-bit register. After the completion of the calibration process the final bits of the SAR register is stored in the N-bit register. The calibration process is controlled by a calibration clock (CLK_{cal}).
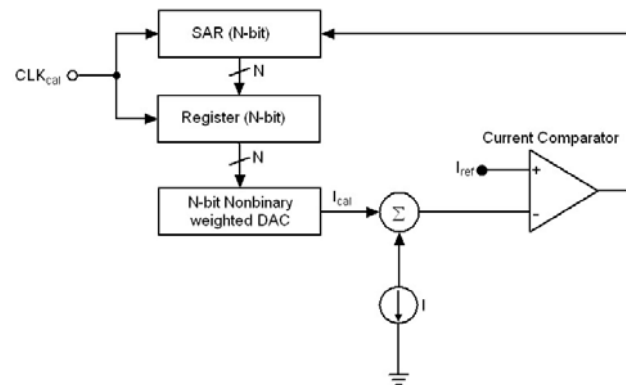

Fig. 4.15. Block diagram of non-binary weighted DAC based calibration loop

In Fig. 4.16 a possible implementation of non-binary weighted calibration DAC is presented [68].   This non-binary weighted DAC has 16 bit input. It is divided into two sub-DACs; a non- binary weighted 8 bit DAC and a binary weighted 8 bit DAC. Both of the sub-DACs have an architecture similar to the conventional binary weighted DAC. Thus the input bit pattern can directly control the current switching.
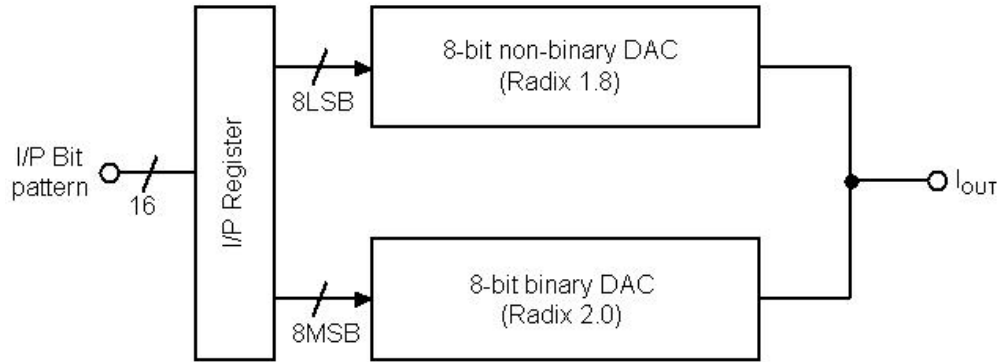
Fig. 4.16. Block diagram of N-bit non-binary weighted calibration DAC

In Fig. 4.17 block diagram of 8-bit non-binary weighted DAC is presented. It is implemented with the radix 1.8. The weighted current sources are implemented with a modified resistive ladder. This resistive ladder has the same architecture like R-2R ladder. But to implement the weighting factor of 1.8 among the current cells the 2R resistance is replaced with a resistance value of 1.8R. The current switches are implemented with conventional CMOS differential pair. At the output of the DAC one of the differential paths is dumped by a resistor and other path is used as the single ended output (I_{OUT}).
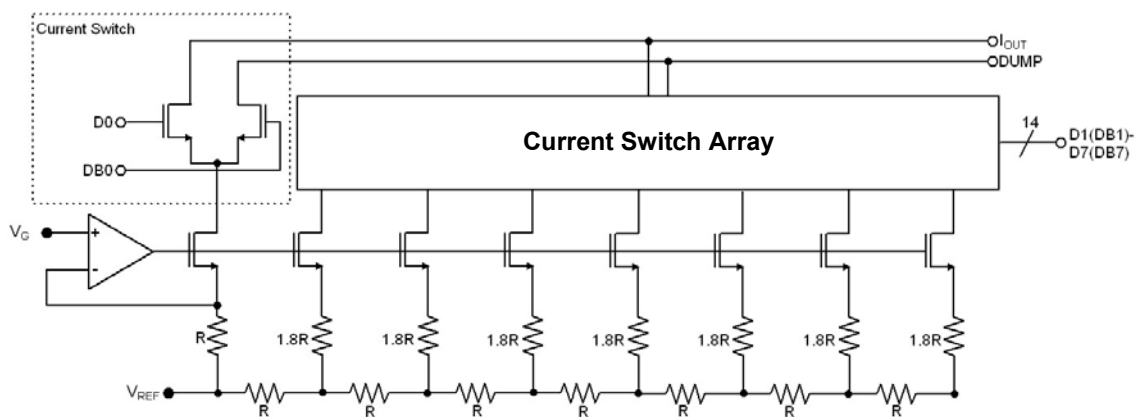

Fig. 4.17. Block diagram of 8-bit non-binary weighted DAC

The 8-bit binary weighted sub-DAC has been implemented in the same fashion of the non-binary weighted sub-DAC. The only difference can be found in the resistive ladder

network. As this sub-DAC is a purely binary weighted so an R-2R ladder network is used unlike R-1.8R ladder network.

The 16-bit non-binary weighted DAC has already been designed in IHP's 0.25µm CMOS technology and the layout is presented in Fig. 4.18. It has ultra low power of 100µW and the area is 0.015mm^2. This non-binary weighted DAC can calibrate up to 0.01% of accuracy.
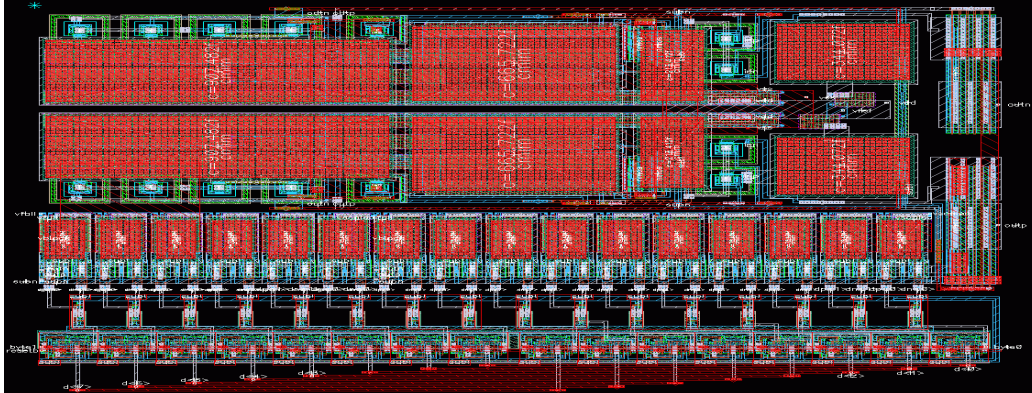
Fig. 4.18. Layout of 16-bit non-binary weighted DAC

The ultra low power and ultra small non-binary weighted DAC can be used to calibrate the individual current cells of binary weighted DAC of a medium resolution (4-6 bit) unary weighted DAC. The main disadvantage of this calibration loop is the exponential increase of number of non-binary weighted DAC with the resolution of input of unary weighted current steering DAC.

4. 5. Conclusions

The upcoming applications in communication systems require high speed medium to high resolution DACs. Current steering architecture is the most suitable candidate to accomplish these applications. The current steering architecture comes in different variants, in which the segmented current steering architecture is most commonly used.

The current steering architecture has a high matching requirement among the current cells, which leads to static INL and DNL error. These performances can be enhanced by the proper sizing of the MOS current sources. On the other hand the main reasons for the deterioration of the dynamic performance are identified as the input code dependent output impedance at the DAC output, the cell dependent switching delay and the current switch nonidealties.

Three different techniques to enhance the performance of the current steering DAC have been discussed. The layout technique is very effective to improve the static performance of the DAC but it comes with large area overhead and complex routing requirement. On the

other hand dynamic element matching technique is very useful to improve both the static and dynamic performances but it is only useful with oversampling DACs. A new non-binary weighted DAC based current cell calibration technique is proposed which is very useful for the high speed DAC and requires very low power and small area.

Chapter 5
Design of Multi-GHz DAC

_____

5. 1. Introduction

The recent growth in the telecommunication market has made the interface between analog and digital parts of the system a critical component. The upcoming application in the multi-gigabit communication systems e. g. radar or satellite communication systems require low to medium resolution (4 to 8-bit) DACs with the multi-GHz sampling rate [44]. On the other hand such high speed DACs can by used for UWB pulse synthesis [45].

To serve these upcoming applications design examples of multi-GHz 4-bit and 8-bit DACs have been presented in this chapter. The 4-bit DAC is implemented with a modified binary current steering architecture. Unlike the binary current steering architecture all of the current cells have the same weightage. The binary weighting operation is implemented with a modified resistive ladder. On the other hand the 8-bit 20GHz DAC is implemented with a modified segmented current steering architecture. 50% segmentation is used. 4LSB bits are converted with R-2R ladder sub-DAC and the rest of the 4-bits are implemented with unary current steering architecture.

This chapter is organized as follows: In section 5.2 the architecture of the 8-bit segmented current steering DAC along with the design of its different sub-blocks are presented. The simulation results of the 8-bit 20GHz segmented current steering DAC has been presented in section 5.3. The measurement results of the 4-bit LSB sub-DAC (presented in Section 5.2.1) of the full 8-bit DAC have been presented in section 5.4. Finally conclusions are drawn in section 5.5.

5 .2. Implementation of High-Speed Segmented Current steering DAC

In chapter 4, architectures of different DACs have been presented. The segmented current steering architecture is found to be the most commonly used DAC architecture. In Fig. 5.1 the block diagram of an 8-bit segmented current steering DAC is presented. Unlike

the conventional segmented current steering DAC (presented in chapter 4) the LSB DAC is implemented with the resistive network. Then the output of the LSB sub-DAC and the MSB sub-DACs are combined to achieve the 8-bit DAC output.
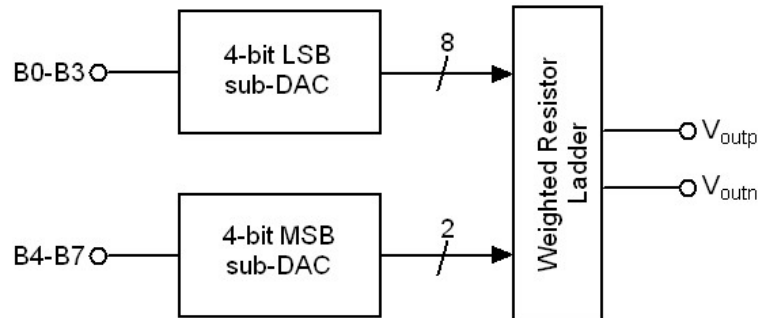

Fig. 5.1. Block diagram of 8-bit modified segmented DAC architecture

The percentage of segmentation is dictated by the static accuracy (INL and DNL) and the area. In the context of multi-GHz DAC design the length of the clock path is a very important issue. With the increasing percentage of segmentation the number of unit current cells of the MSB sub-DAC increases exponentially so does the length of the clock path. As a result the delays among the current cells become unequal, which increases the spurious free dynamic range (SFDR, see section 5.3.2.2). A compromise is found in 50% of segmentation for the 8-bit DAC implementation. In the following sub-sections the implementation of different sub-blocks of the segmented current steering DAC has been presented.

5. 2. 1. Design of 4-bit LSB Sub-DAC

The LSB sub-DAC, which has weighted resistor architecture, can be considered as a special kind of binary weighted DAC, where the current weighting function is implemented by the resistive ladder network. Four unary weighted current sources with the current weightage I_{LSB} are connected with ladder network. A simplified block diagram of this LSB sub-DAC is presented in Fig. 5.2.

The input bit patterns are stored in the 4-bit input register. Then these input bits are delayed by a full clock cycle to equalize the delay of the unary and the binary weighted sub-DACs. The outputs of this delay matching register are synchronized with the input clock edge by the retiming D flip-flops (DFF), which control the switching of unit current cells. Binary weighting operation of these unit current cells is accomplished by the resistive ladder network. In the following sub-sections the design of different sub-blocks for the 4-bit LSB sub-DAC are presented.
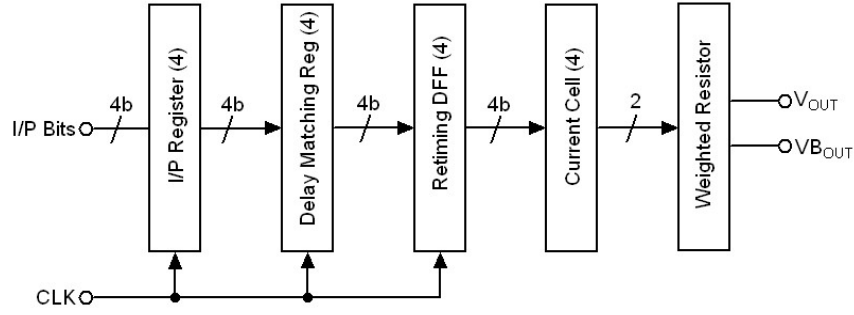
Fig. 5.2. Block diagram of LSB DAC

5. 2. 1. 1. Design of Input and Delay Matching Register

In Fig.5.2 the block diagram of the 4-bit LSB sub-DAC has been presented. The four input LSBs (B0-B3) are stored in parallel out the input register. Then these input bit pattern (B0-B3) is delayed by one full clock cycle to achieve concurrency in the output of LSB and MSB sub-DACs. This delay is implemented with another 4-bit register known as delay matching register, which essentially has the same architecture like input register.

A conventional ECL master slave DFF is used as the building block of this 4-bit register. A simplified block diagram of this ECL master slave DFF (MSDFF) is presented in Fig. 5.3. The MSDFF consists of two identical ECL D-latches. The input differential signals (D, DB) are connected with the master latch and the slave latch provides the differential output (Q, QB). These cascaded master and slave ECL D-latches are controlled by a differential input clock (CLK, CLKB). The master and slave latches work in the time interleaved fashion. This is accomplished by twisting the differential clock at the clock input of the slave D-latch.

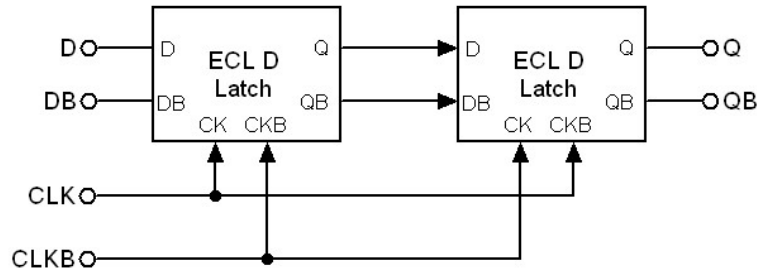Fig. 5.3. Block diagram of ECL master slave DFF

A simplified schematic diagram of the ECL D-latch is presented in Fig. 5.4. It is implemented with the conventional ECL D-latch architecture [69]. All of the transistors used in this design are having same emitter size. To achieve higher speed only minimum emitter size transistors are used. The input differential signals are composed of D and DB whereas

the differential clock signals are constructed with CK and CKB. The differential clock signal (CK, CKB) has a common mode level, which is lower than that of the input differential signal (D, DB) by the amount of base-emitter voltage of transistor Q1.

The operation of this ECL D-latch can be divided into two phases. In the first phase when differential clock is high (i.e. CK is higher than CKB) tail current I_1 flows through transistor Q5. In this phase Q1 and Q2 appear as the input differential pair and the output differential nodes (Q, QB) are charged according to the input differential signal (D, DB). The ECL latch goes into the next phase when differential clock signal goes low (i.e. CK is lower than CKB). In this phase the tail current I_1 is switched from transistor Q5 to Q6. Thus the input differential pair (Q1 and Q2) becomes inactive in this phase whereas the other differential pair (constructed with Q3 and Q4) starts to conduct. The bases of the differential pair Q3 and Q4 are connected with the differential output nodes Q and QB, which appear as a regenerative stage and sustain the differential outputs unchanged during this phase.

In this current design 2.5V of power supply is used with a differential logic swing of 600mV$_{PP}$. The common mode voltage of the differential inputs (D, DB) and that of the differential clock are 2.35V and 1.45V respectively. It shows a typical delay of 12pS with the tail current (I_1) of 3mA.
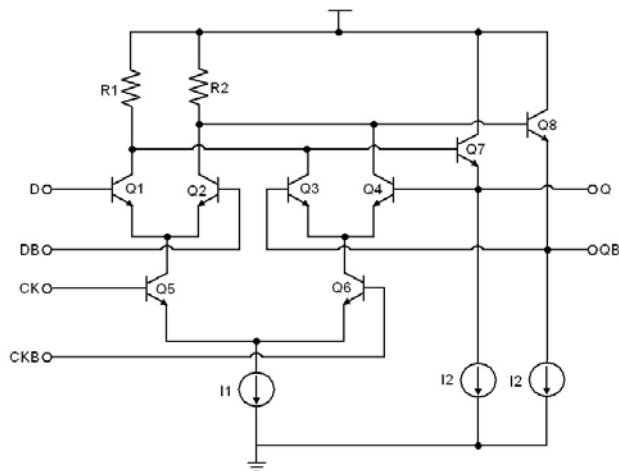
Fig. 5.4. Simplified schematic of ECL D-latch

5. 2. 1. 2. Design of Unit Current Cell

In Fig. 5.5 the schematic of a simple unit current cell has been presented. This unit current cell has two main components, the tail current source and the current switch. The current switch has been implemented with an HBT emitter-coupled pair. An improved unit

current source, which provides higher output impedance, is presented in Fig. 5.6. Unlike the simple unit current source it has pair of cascode devices on top of the main differential current switch.
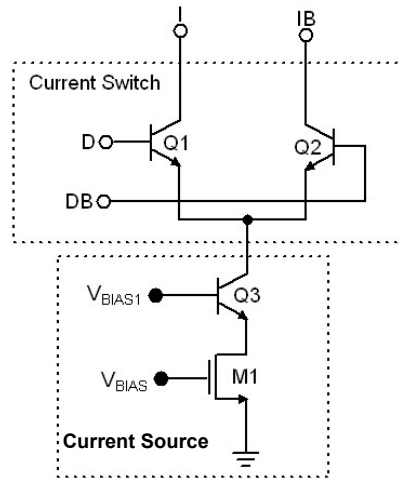
Fig. 5.5. Simplified schematic of unit current cell

In section 5.3.2.1 it has been presented that the output impedance of the current source has a direct impact upon the dynamic performance of the DAC. According to equation 5.15 the most suitable way to enhance the dynamic performance is to increase the output impedance of the current source. Due to this reason cascode current mirrors are used as the current source in both of the unit current cells. In Fig. 5.6 conventional HBT cascode current
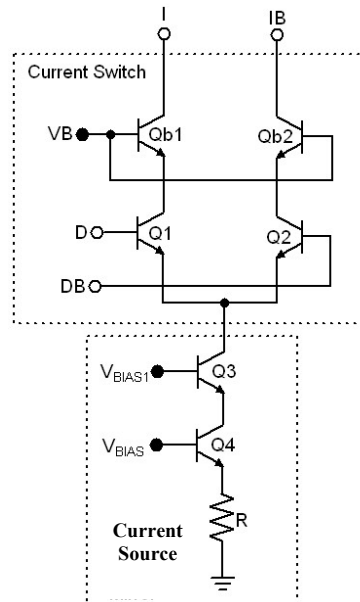
Fig. 5.6. Schematic of improved unit current cell

mirror is used. On the other hand in Fig. 5.5 an nMOS transistor is used as the main current source. In both the current sources a common base HBT is used as the cascode device. By optimizing the area of the nMOS transistors a specified percentage of matching accuracy can be achieved. But the output impedance of the nMOS transistor decreases rapidly with the increasing frequency. As a result of it this current source is not very useful for high frequency high resolution applications. In section 5. 4, it has been shown that this current source can be used for 4-bit of resolution with 30GHz of sampling rate. But with the increasing resolution the number of current cells are increased and the combined output impedance of the parallel current sources is reduced. In this case the HBT cascode current sources come as the better choice. Hence for the 8-bit, 20GHz DAC design the HBT cascode current source is used (as shown in Fig. 5.6). For these multi-GHz applications sometimes even the HBT current source does not provide sufficiently high output impedance, which reduces the dynamic performances of the DAC. In this case another pair of cascade transistors can be used on top the differential current switch (see Fig. 5.6).

The current switch is implemented with a simple differential pair (see Fig. 5.5). The main errors associated with the current switches are the clock feedthrough and the charge injection. For the HBT current switches the charge injection is not an issue because of the fact that the base excess charge recombines in the intrinsic base region. But the feed through of the input signal to the output is a major problem in the high speed applications. The base-collector parasitic capacitance couples a considerable amount of input signal to the output. Reducing the input signal swing to control the current switching can reduce the feedthrough. In this particular application $600mV_{pp}$ differential signal is used.

5. 2. 1. 3. Design of Retiming DFF

In the section 4.3.2.2 it is presented that delay spread among the current cells results in higher 2nd order harmonic. The retiming DFF decides the precise switching instances of the current cells synchronized to the clock rising or falling edge. In Fig. 5.7 the block diagram of a retiming DFF is presented. The core of this DFF is implemented with the conventional ECL DFF as discussed in section 5.2.1.1. The outputs of this DFF are synchronized with the falling edge of the input differential clock (CLK and CLKB). In Fig. 5.8 a typical output of the ECL DFF is presented, where the outputs are synchronized with the falling edge of the clock.
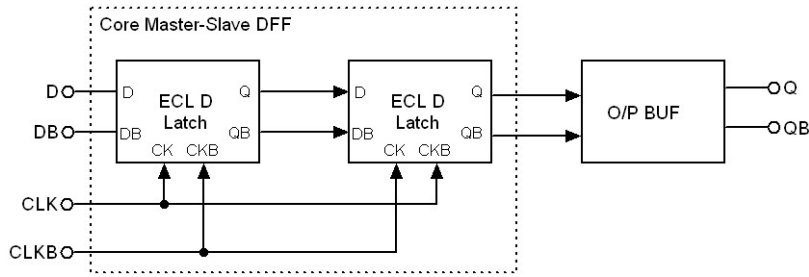
Fig. 5. 7. Block diagram of retiming DFF

Depending upon the intrinsic time constant and the output load the DFF shows a finite delay and rise or fall time. In addition to that the differential outputs show high frequency glitches at the rising edge of the clock (CLK). This glitch is caused due to the switching of the current from one differential pair of the D latch to the other (see Fig.5.8). This high frequency glitch at the rising edge of the CLK, directly couples to the output of the current cell and causes a current glitch.

To overcome this problem a high gain high bandwidth output buffer is used at the output of the core DFF. This output buffer works as limiting amplifier and reduces the output glitch of the retiming DFF. A simple two stage differential amplifier is used as this output buffer.
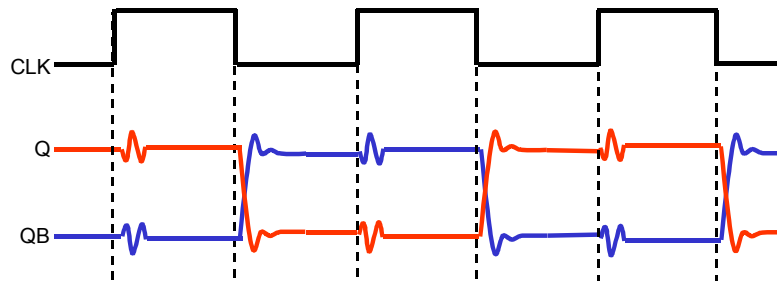
Fig. 5. 8. Output waveform of an unbuffered DFF

5. 2. 1. 4. Design of Weighted Resistor Network

Unlike the conventional binary weighted current steering DAC, the LSB sub-DAC is having the same weight for all of the current cells and the binary weighting operation is accomplished by the weighted resistive ladder. The most commonly used resistive ladder is the R-2R network as presented in section 5.2.4. The schematic of this R-2R ladder is presented once again in Fig. 5.9 for the 4-bit LSB sub-DAC. The main advantage of this ladder network is its symmetrical and modular structure, which provides a great advantage in
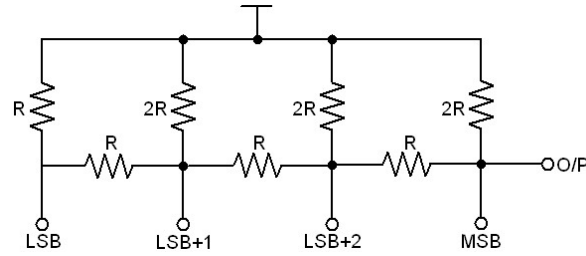
Fig. 5. 9. R-2R Ladder network for 4-bit DAC

high resolution DAC design. This R-2R ladder network can be designed to match directly with the external load resistance of 50Ω. This R-2R resistive ladder shows different delays to the output for the different current cells. Generally this delay variation is small compared to the sampling time period. But for multi-GHz DAC design this can deteriorate the dynamic performances for high resolution DACs.

In Fig. 5.10 another variant of weighting resistor network has been presented [70] for a 4-bit DAC. The output impedance of this resistor network is 8R, which is directly matched with the external 50Ω load. Compensation resistors are used to have same resistive load at the output of all current cells. Unlike the R-2R ladder all of the input nodes of this ladder network has the same potential thus the current switching dynamic is very much identical for this kind of resistive ladder. The main disadvantage of this resistive network is its asymmetric architecture. For the 50Ω output load unit resistance R is 5.25Ω. Such a precise low resistance fabrication is difficult in sub-micron technologies. Moreover the parasitic interconnect resistances also reduces the matching accuracy. Thus in spite of having better dynamic performance than the R-2R ladder, this resistive ladder is not very suitable for the high resolution DACs. As a result, this resistive ladder is used to implement the 4-bit standalone DAC (presented in section 5.4) whereas the R-2R ladder network is used for the design of 8-bit segmented current steering DAC.
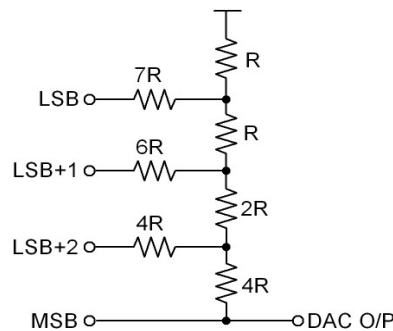

Fig. 5.10. Schematic of modified weighted resistor network

97

5. 2. 2. Implementation of 4-bit MSB Sub-DAC

The 4-bit MSB sub-DAC has been implemented with the unary weighted current steering architecture. The detail description of this unary weighted current steering architecture has been presented in section 5.2.2. A simplified block diagram of the 4-bit unary weighted MSB sub-DAC is presented in Fig. 5.11. The four MSBs (B4-B7) are stored in the input register. This input register has the same architecture as discussed in section 5.2.1.1.
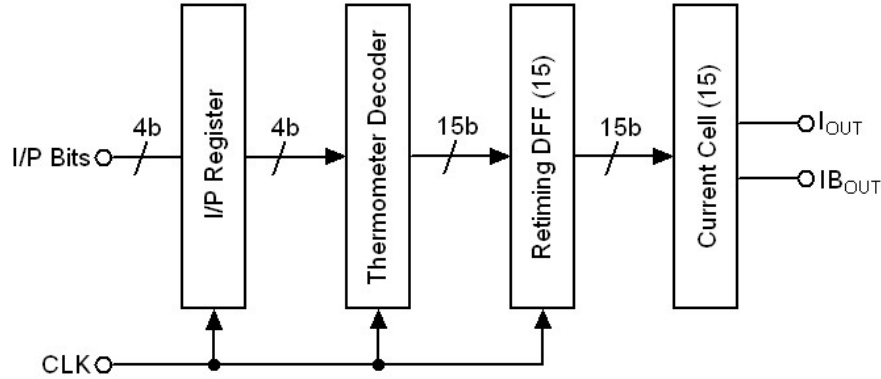

Fig. 5.11. Block diagram of 4-bit MSB Sub-DAC

The binary coded inputs are then converted into thermometer code by the means of thermometer decoder. This thermometer decoder is a combinational circuit. Thus the thermometer-coded outputs could have different delays. To make these outputs concurrent with the input clock (CLK) the retiming DFF array is used. The retiming DFF directly controls the current cell unit. These retiming DFF and the current cell units have the same architecture described in sections 5.2.1.3 and 5.2.1.2 respectively.

As discussed in the chapter 4.2.2 the design of high-speed thermometer decoder comes as a bottleneck in the unary weighted DAC particularly when the conversion speed is in the range of few tens of gigahertz. In the later section a new technique of the design of thermometer decoder is presented, which is particularly adapted for high-speed application.

5. 2. 2. 1. Design of High-speed Thermometer Decoder

In Fig.5.12 the block diagram of a commonly used thermometer decoder [59] for unary weighted DAC is presented. In this implementation the N-bit binary to thermometer decoding operation is accomplished in two steps. The P no. of LSB bits are connected to the column decoder and Q no. of MSB bits are connected to the row decoder such that; N=P+Q. The P-bit column decoder provides (2^P-1) thermometer coded outputs ranges from C[0] to C[2^P-1]
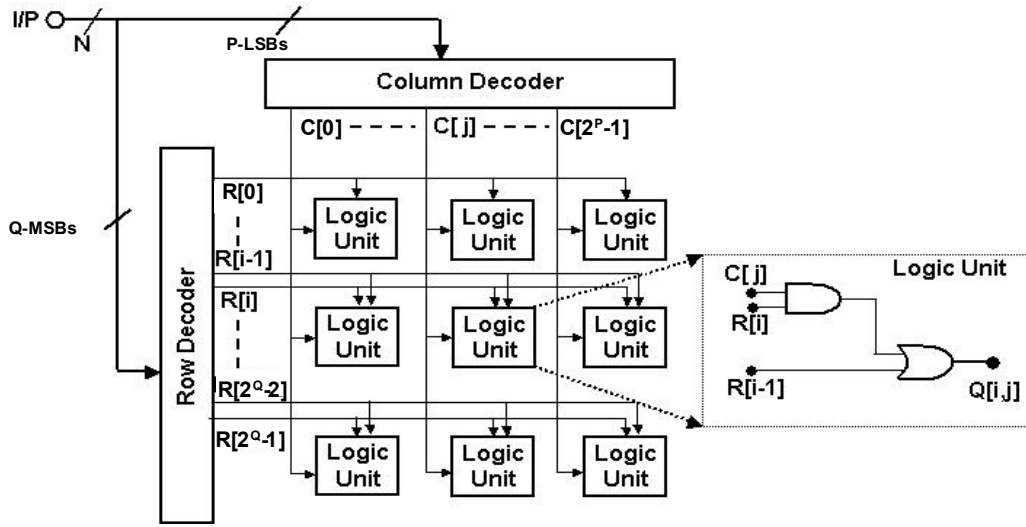
Fig. 5.12. Conventional binary to thermometer decoder

similarly the outputs of row decoder ranges from R[0] to C[2^Q-1]. C[0] and R[0] are always logic high for any input. The outputs of the Row and Column decoders are combined in local combinational logic unit to achieve the required thermometer coded outputs. In Fig. 5.12 the outputs of the thermometer decoder are arranged in a two dimensional matrix. For any output of Q[i.j] the combinational logic can be expressed as,

$$Q[i, j] = R[i] \bullet C[j] + R[i-1] \tag{5.1}$$

where, R[i] and R[i-1] are the ith and (i-1)[th] outputs of the Row decoder respectively. C[j] is the jth output of the Column decoder.

For a particular application in the 4-bit binary to thermometer decoder this afore mentioned approach can be implemented in the following fashion. The 4-bit input is connected with two 2-input Row and Column decoders.

For a 2-bit column decoder the input bits are defined as B0 and B1 whereas the outputs are defined as C[0], C[1], C[2], C[3]. As mentioned in earlier C[0] is always logic high. The rest of outputs of the column decoder can be expressed by the following,

$$\begin{aligned}
C[1] &= B1 + B0 \\
C[2] &= B1 \\
C[3] &= B1 \bullet B0
\end{aligned} \tag{5.2}$$

Similarly the outputs of the row decoder can be defined as R[0], R[1], R[2] and R[3] for the input bits B2 and B3. R[0] is logic high. The relation among the inputs and the rest of the outputs can be represented by the Equation 5.3.

$$R[1] = B3 + B2$$
$$R[2] = B3$$
$$R[3] = B3 \bullet B2$$

(5.3)

It can be observed from Equation 5.2 and Equation 5.3 that the decoder outputs R[3] and C[3] have the highest delay. The outputs of the Row and Column decoders are combined in a combinational logic block according Equation. 5.1 to get the desired thermometer coded output.

The maximum possible delay from the input to the output of the thermometer decoder has been shown in Fig. 5.13. Here it is assumed that any input and its inverted signals are available concurrently. This assumption holds for the ECL or CML logic gates because those gates provide differential outputs. Output Q[3,3] has the maximum delay from the input to the output. This delay is the summation of two AND gates and one OR gates. The conventional ECL AND and OR gates [69] are designed and simulated in IHP's 0.25μm SG25H1 technology [34]. For these designs 3.0mA tail current is used for both of the gates with 2.5V of power supply. In simulation these ECL AND and OR gates show typical delays of 12pS and 10pS respectively. Thus the maximum delay from the input to the output of 4-bit binary to thermometer decoder will be 34pS.

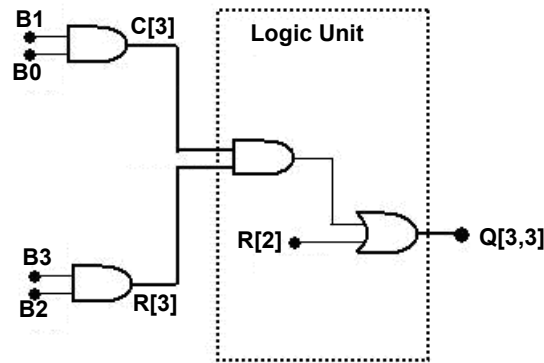Fig. 5.13. Longest delay path from the input to the output

In the segmented current steering DAC the thermometer decoder lies in between the input register and the retiming DFF (as shown in Fig. 5.11). Both input register and the retiming DFF consist of ECL MSDFFs and are controlled by the same input clock (CLK). Thus the output of the thermometer decoder has to be settled down within half of the input clock

100

period. Thus according to the given design example the thermometer decoder with the longest delay of 34pS can work up to 14.7GHz.

One of the most commonly used techniques to enhance the speed of the combinational circuit is to break the complex logic operation into simple parts and introduce registers in between these combinational blocks. In high speed logic design this technique is not very useful as the power dissipation and the area are increased. In Fig. 5.14 the block diagram of an improved 4-bit binary to thermometer decoder architecture [71] along with the input and output interfaces are presented. Unlike the conventional thermometer decoder here the main decoding operation is done by a bipolar ROM and the address decoding for this ROM is accomplished with the binary decoder.
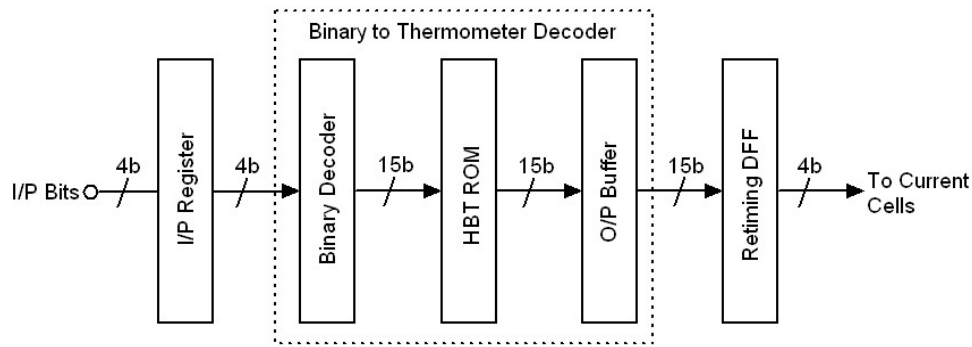
Fig. 5.14. Block diagram of improved 4-bit binary to thermometer decoder

The 4-bit binary decoding needs complex combinational operation. This imposes speed limitation due to the gate delay of the combinational logic. In this proposed architecture a new method is adopted to design high-speed combinational logic. The principle of this implementation was proposed in [65]. It has been shown that the wired OR/NOR function can be merged with the conventional ECL D-latch. The block diagram of such an N-input OR/NOR DFF is presented in Fig. 5.15.
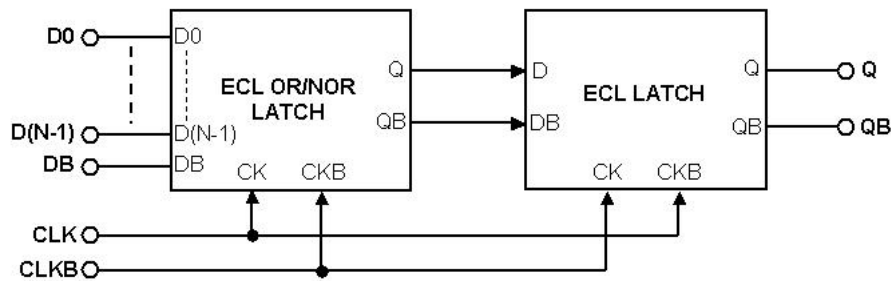
Fig. 5.15. Block diagram of OR/NOR ECL DFF

This OR/NOR DFF has the similar architecture to the MSDFF. The only difference can be found in the master latch. In this present OR/NOR ECL DFF the master latch has been changed with an N-input ECL OR/NOR DFF. This latch works with single ended inputs (D0-D(N-1)) and DB is connected to a DC voltage, which defines the logic threshold level for the inputs. At the output node Q it provides the logical OR output of the N-inputs whereas output node QB corresponds to logical NOR of the inputs. The slave latch has the same architecture as shown in Fig. 5.4.

The schematic diagram of the 4-input ECL OR/NOR DFF is presented in Fig. 5.16. Unlike the conventional ECL D-latch it has 4 transistors (Q1-Q4) in parallel, which perform the wired OR/NOR function for the inputs D0 to D3. As mentioned earlier this OR/NOR latch works with single ended inputs so it has AC gain 3dB less than that of the conventional ECL latch. Although it does not reduce the output logic swing if the single ended logic swings are high enough. In addition to that the regenerative transistor pair Q6 and Q7 achieves the full logic swing in the hold mode (i.e. when CK is low compare to CKB)
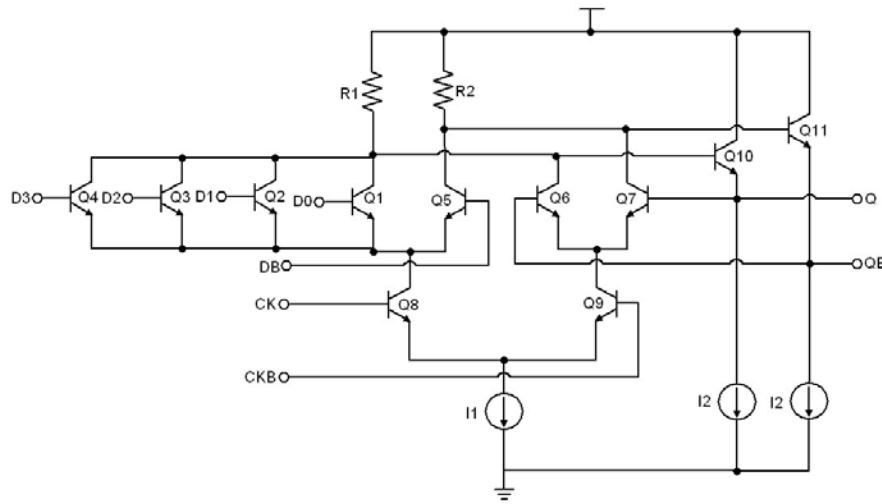
Fig. 5. 16. Schematic of 4-input OR/NOR DFF

With increasing number of inputs the parasitic load capacitance at the output node QB gets increased, which increases the delay at the output. In Fig. 5.17 the simulated output delay of the ECL OR/NOR D-latch is plotted with the increasing number of inputs. This ECL OR/NOR D-latch is implemented with IHP's 0.25μm SG25H1 technology with 2.5V of power supply. The tail current (I_1) is 3mA with 20fF of load at each output. For the inputs 300mV$_{PP}$ single ended logic swing is used with the DC level of 2.35V. A 600mV$_{PP}$ differential input clock is used and in this plot the clock frequency is 20GHz. The worst case

delay occurs when only one out of the N-inputs is high. Under this circumstance one input transistor is charging or discharging the output node QB (see Fig. 5.16).

From Fig. 5.17 it can be observed that the delay of the OR/NOR D-latch increases linearly with the increasing number of inputs. For the incremental delay measurement 1-input ECL OR/NOR D-latch is considered to be the reference, then the number of inputs are gradually increased to 5. The difference between the delays of (i-1)-input OR/NOR D-latch with that of i- input OR/NOR D-latch is defined as the incremental delay for the i-input OR/NOR D-latch. This incremental delay is almost constant and the average incremental delay is 0.98pS per input increment. A 4-input OR/NOR D-latch shows typical absolute delay of 14.2pS.
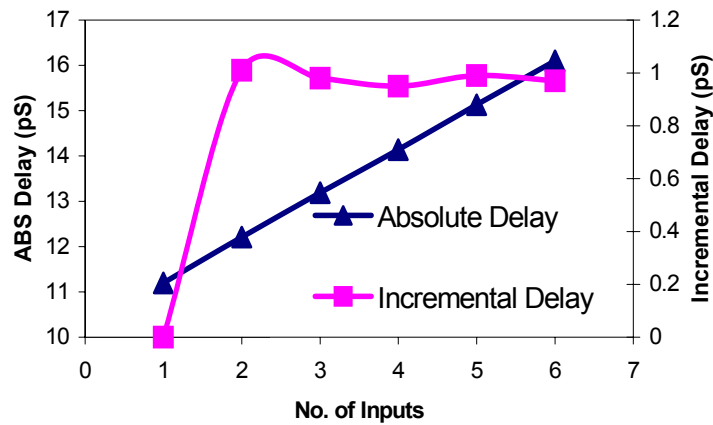
Fig. 5.17. Plot of absolute and incremental delay with increasing no. of inputs for OR/NOR D-latch

This 4-input OR/NOR D-latch can be used as the master D-latch to implement 4-input master slave OR/NOR DFF as shown in Fig. 5.15. In the context of the 4-bit binary decoder design this 4-input master slave DFF can be used as the building blocks. In this design the single ended inputs and outputs are used. Assuming the 4-bit binary decoder has the inputs B0-B3 and the outputs, y[0] to y[15], the relation among the inputs and the output can be expressed by the following equations 5.4,

As it is shown in Fig. 5.1. the input bits of the thermometer decoder are provided by the input register and the input bits (B0-B3) and their complementary bits ($\overline{B0}$ - $\overline{B3}$) are readily available. By connecting the 4-input OR/NOR master slave DFFs according to the Equation 5.4 the binary decoding operation is accomplished.

$$y[0] = \overline{\overline{\overline{B3} + \overline{B2} + \overline{B1} + \overline{B0}}}$$
$$y[1] = \overline{\overline{\overline{B3} + \overline{B2} + \overline{B1}} + B0}$$
$$y[2] = \overline{\overline{\overline{B3} + \overline{B2}} + B1 + \overline{B0}}$$
$$\vdots$$

(5.4)

$$\vdots$$
$$y[14] = \overline{\overline{B3} + B2 + B1 + \overline{B0}}$$
$$y[15] = \overline{B3 + B2 + B1 + B0}$$

5. 2. 2. 2. Design of HBT ROM
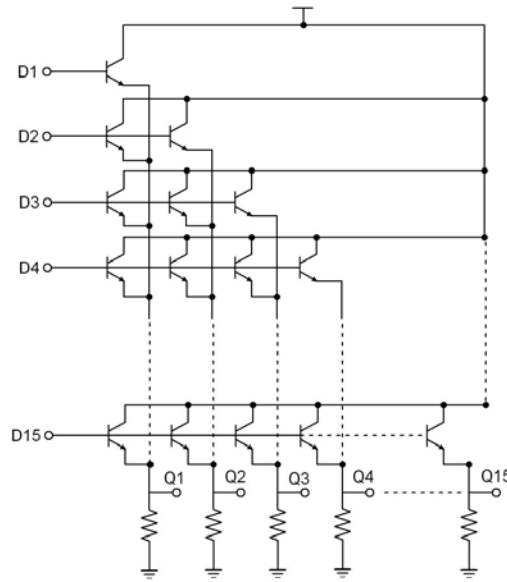
Fig. 5.18. Simplified schematic of pseudo differential ROM

The ROM accomplishes the main binary to thermometer decoding operation. The simplified schematic diagram of HBT ROM is presented in Fig. 5.18 [71]. This ROM has pseudo differential architecture and only one of the pseudo differential parts is shown in Fig. 5.18. The ROM performs wired OR logical operation and can be expressed as,

$$Q1 = D1 + D2 + D3 + \cdots\cdots + D15$$
$$Q2 = D2 + D3 + D4 + \cdots\cdots + D15$$
$$Q3 = D3 + D4 + D5 + \cdots\cdots + D15$$
$$\vdots$$

(5.5)

$$\vdots$$
$$Q14 = D14 + D15$$
$$Q15 = D15$$

where, D1, D2, D3, ……., D15 are the input of the ROM generated by the binary decoder and Q1, Q2, Q3,…….., Q15 are one of the pseudo differential output of the ROM.

As presented in equation 5.5 the logical OR function is implemented with the parallel combination of emitter followers. For any particular input data pattern only of the output of the binary decoder goes high and all other outputs are low. A worst case scenario occurs when the input D1 goes high and rest of the inputs (D2, D3, …., D15) are low. In this case only one of the emitter followers tries to pull the output voltage (Q1) high and rest of the fourteen emitter followers push the output voltage down. With a sufficiently high input voltage (D1) the output goes high but the logic swing get reduced. As a solution of it the logic levels are restored with single differential stage output buffer.

5. 2. 3. Design of 8-bit Segmented Current Steering DAC
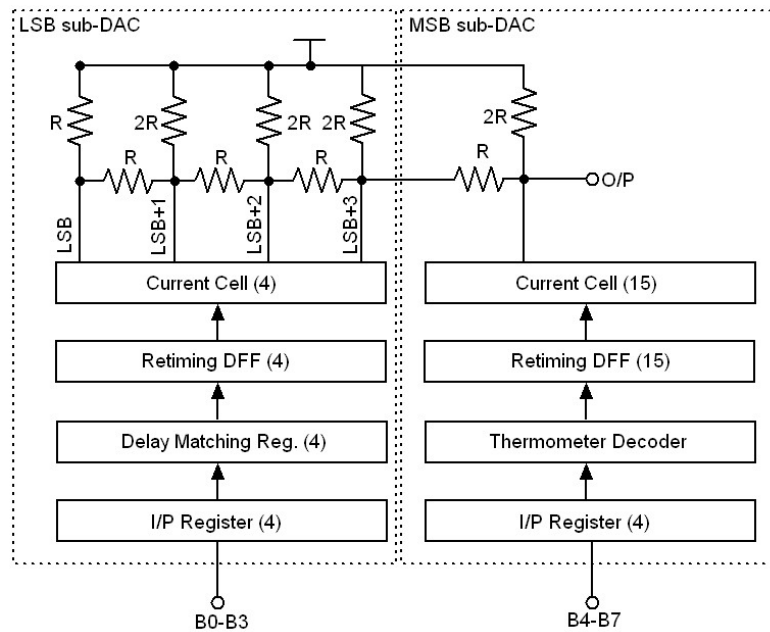
Fig. 5.19. Block diagram of 8-bit segmented current steering DAC

In Fig. 5.19 the block diagram of an 8-bit segmented current steering DAC is presented. Unlike the conventional segmented current steering DAC (presented in chapter 5) the LSB DAC is implemented with the R-2R ladder network. Then the output of the LSB sub-DAC and the MSB sub-DACs are combined to achieve the 8-bit DAC output.

The R-2R based LSB sub-DAC has number of advantages over the binary weighted DAC. Unlike the binary weighted architecture in R-2R DAC architecture all of the current cells have the same weight, which improves the matching among those current cells. The

105

current switch dynamics are very similar in this architecture as all of the switches work with the same weight of current source and the output impedance is always constant [72]. Particularly when the sampling speed is in multi-GHz range this architecture comes with a great advantage in terms of output matching. The R-2R ladder can directly be matched with the external 50Ω load.

The floorplan of the DAC plays a critical role for the static and dynamic performance. The floorplan is directly driven by the requirements of the dynamic performances. In section 4.3.3 it has been presented that the variation of switching delay causes higher second order harmonic. A commonly used technique to reduce the clock delay has been presented in Fig. 5.20 [4]. In this technique a tree-like clock and output lines are used to compensate the
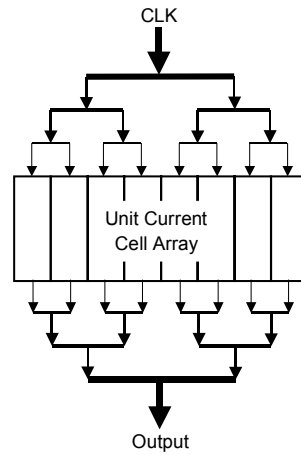
Fig. 5. 20. Tree-like clock and output routing

output clock and output delays of different current cells. But unfortunately this type of clock and output routing requires longer path length. An alternative to this routing technique is presented in Fig. 5.21. In this routing technique the current cells are placed in an array (one dimentional). The clock signal and the output signal taps are directly connected with the respective signal paths. These long clock and output signal paths are implemented with 50Ω microstrip transmission lines.
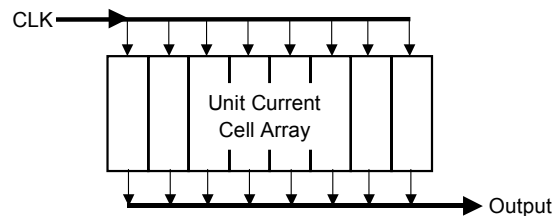
Fig. 5. 21. Delay compensated clock and output routing

5. 3. Simulation Results of the 8-bit Segmented Current steering DAC
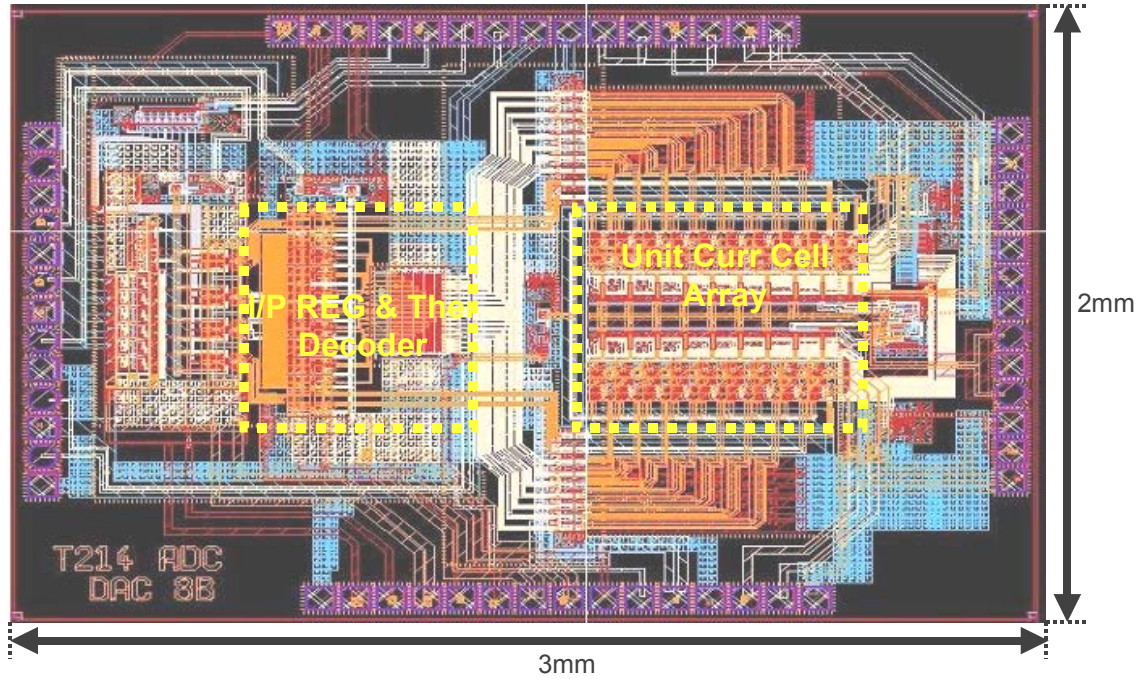
Fig. 5.22. Layout of 8-bit segmented current steering DAC

The 8-bit segmented current steering DAC has been designed with IHP's 0.25µm SG25H1 BiCMOS process with three thin metal layers and two thick top metal layers. The HBTs have f_t and f_{max} of 190GHz. Additionally this technology provides metal-insulator-metal (MIM) capacitors and poly-silicon resistors. The complete layout of the 8-bit DAC is presented in Fig. 5.22. The total chip area is 6mm^2. This chip has two different power supplies. The main analog part i.e. the unit current cells and the R-2R ladder network works with 4.5V, whereas the digital parts of the DAC works with 2.5V of power supply. The full chip consumes 2.5W of power.

In high speed design the passive interconnects play a critical role for the static and dynamic behavior. In the DAC design relatively long interconnects (e. g. data, clock, output lines) are implemented with 50Ω microstrip transmission lines. These microstrip transmission lines are simulated in 2.5D electromagnetic simulator (ADS Momentum) and equivalent π-models for there passive transmission lines are generated. In the simulation these models of the transmission lines are incorporated to have the more realistic results. In Fig. 5.23 a portion of the transfer characteristics of the DAC is presented for the input clock rate of 20GHz. In this simulation a digital ramp function is used. The input bit pattern is started from "00000000" and in each step the input is incremented by "00010001". Thus in each step

transition one additional current cell from the MSB sub-DAC is switched to the output and additionally the input for LSB sub-DAC is incremented by 1. These special input patterns are used to reduce the simulation time of the DAC. In Fig. 5.23a the single-ended pair of outputs of the DAC (OP+ and OP-) have been presented whereas the difference between these outputs (Diff O/P) is plotted in Fig. 5.23b. For the full scale transition (i. e. "00000000"→ "11111111" or "11111111" →. "00000000") the DAC output shows rise and fall time of 18.3pS, which shows that the DAC can work up to the Nyquist bandwidth (for 20GHz of clock input the Nyquuist banwidth is 10GHz)
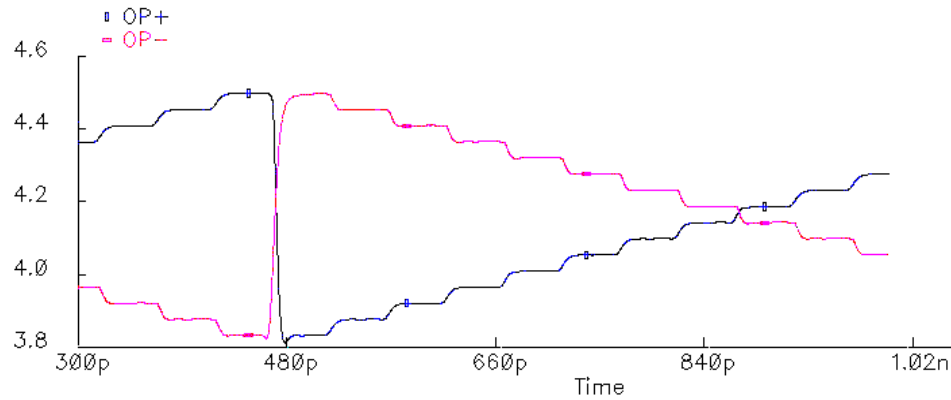
(a)
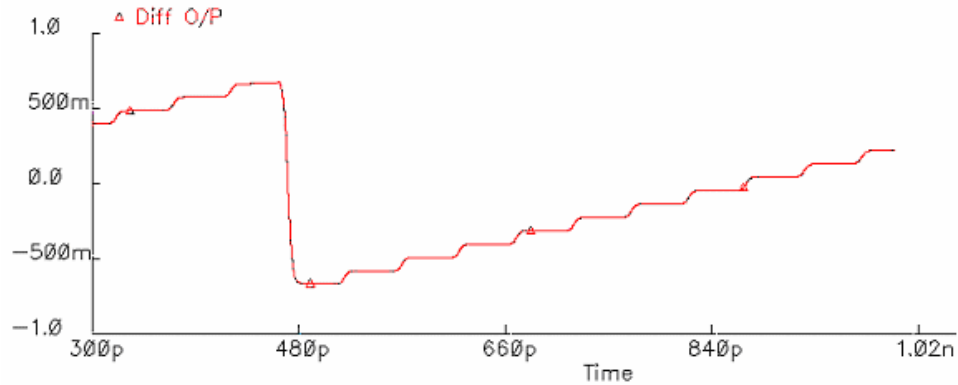
(b)

Fig. 5.23. (a) Single-ended outputs, (b) Differential output of the DAC for digital ramp input

For the dynamic performance analysis the input bit patterns for the DAC are generated with an ideal 8-bit ADC. The digital input is converted back into analog signal. Such a reconstructed output of the DAC is presented in Fig. 5.24. In this plot a 9GHz sinusoidal is used (F_{in}) and the sampling rate of the DAC (F_s) is 20GHz.
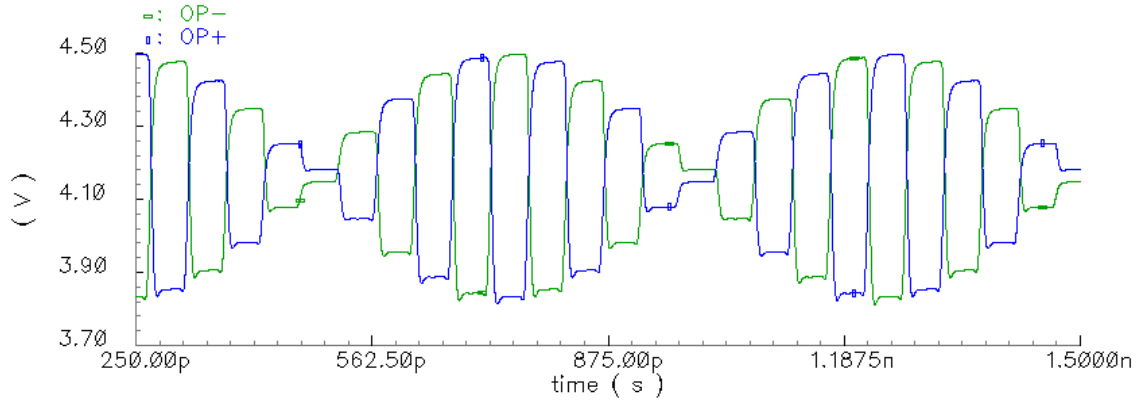
(a)
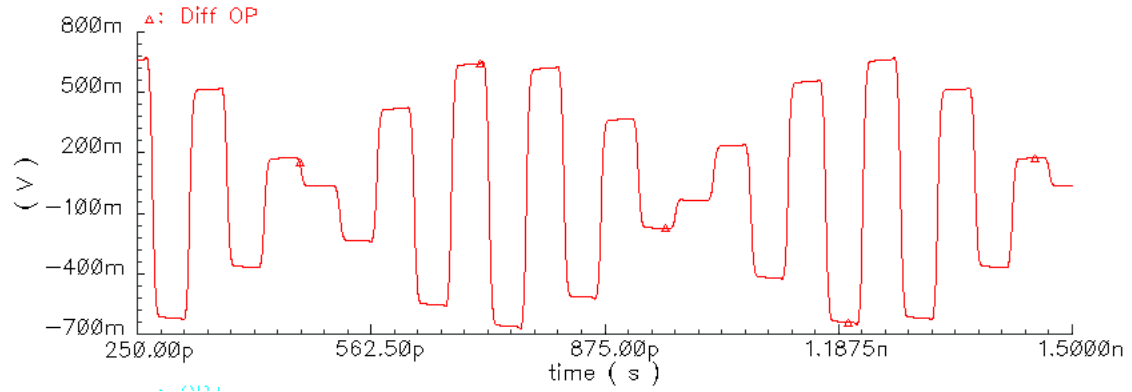
(b)

Fig. 5. 24. (a) single-ended, (b) differential output signal of the DAC for F_{in}=9GHz, F_s=20GHz

The accuracy of this 8-bit DAC has been estimated in the frequency domain. As in the simulation the delay between the differential signals is always equal so the second order harmonic will show very low amplitude. Hence, for the accuracy calculation this second order harmonic is neglected and the difference fundamental and the third order harmonic is approximated as the total harmonic distortion (THD). The THD is used as the figure of merit to calculate the accuracy of the DAC in terms of effective number of bits (ENOB). The relation between the ENOB and THD is presented in chapter 2 and it is once again presented here,

$$ENOB = \frac{THD(dB) - 1.76}{6.02} \qquad (5.6)$$

where, THD is expressed in dB.

In Fig. 5.25 output spectrum of the 8-bit DAC has been presented. In this spectrum a full-scale 9GHz sinusoidal digital input is used and the input clock rate (F_s) is 20GHz. The difference fundamental (9GHz) and the third order harmonic (27GHz) is –48.9dBc. This
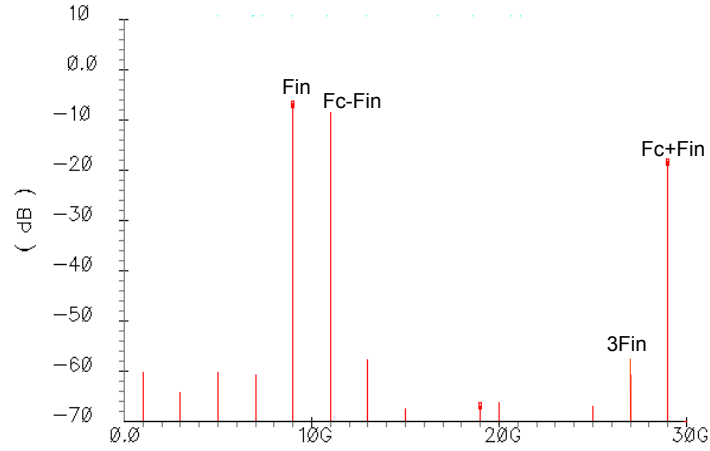
Fig. 5. 25. Output spectrum of the 8-bit DAC for F_{in}=9GHz and F_s=20GHz

difference between the fundamental and the third order harmonic is approximated as the THD. According to equation 5.6 this THD corresponds to 7.83ENOB.

In Fig. 5.26 the amplitude of fundamental and the third order harmonics have been presented for different input frequencies. In this simulation full scale sinusoidal digital patterns are used as the input and the clock frequency (F_c) is 20GHz. From the Fig. 5.26, it can be seen that the full input frequency range the fundamental frequency has almost the flat amplitude values and so does the third order component. Thus the DAC has almost constant linearity for the frequency range from 4GHz to 9GHz. It shows the lowest ENOB for the 9GHz input signal and it is 7.83-bit. In table 5.1 the summarized simulation results of the 8-bit 20GHz DAC has been presented.
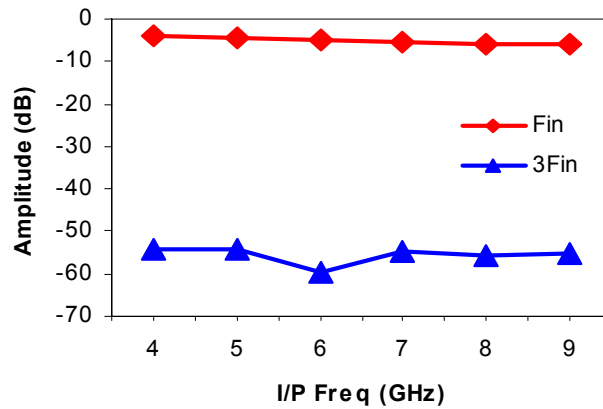
Fig. 5. 26. Fundamental and 3rd order frequency components for different input frequencies

Table 5.1. Summarized simulation results for 8-bit 20GHz DAC

Process	IHP's 0.25μm SiGe BiCMOS SG25H1
Resolution	7.83 bit
Conversion rate	20 GHz
Output resolution bandwidth	9GHz
Supply voltage	2.5V/3.5 V
Power dissipation	2.5W
Die area with pads	6 mm^2

5. 4. Measurement Results of 4-bit Modified Binary Weighted DAC

The 4-bit LSB sub-DAC (as presented in section 5.2.1) has designed separately and fabricated. The simple current cell unit as shown in Fig. 5.5 is used for this implementation. The schematic of the resistive ladder is presented in Fig. 5.10.   This 4-bit DAC can work up to 30 GHz of input clock rate. This 4-bit DAC has been implemented in IHP's 0.25 μm 190 GHz BiCMOS SG25H1 technology [34]. The DAC was designed in a test chip together with some other blocks. The chip micrograph of the 4b-DAC section of the test chip is shown in Fig. 5.27. The core area is 0.70 mm^2. The full DAC works with 3.5V power supply. It dissipates 455mW of power, including an on-chip clock driver. The output buffer of the retiming DFF and the clock driver consume almost 70% of the total power dissipation
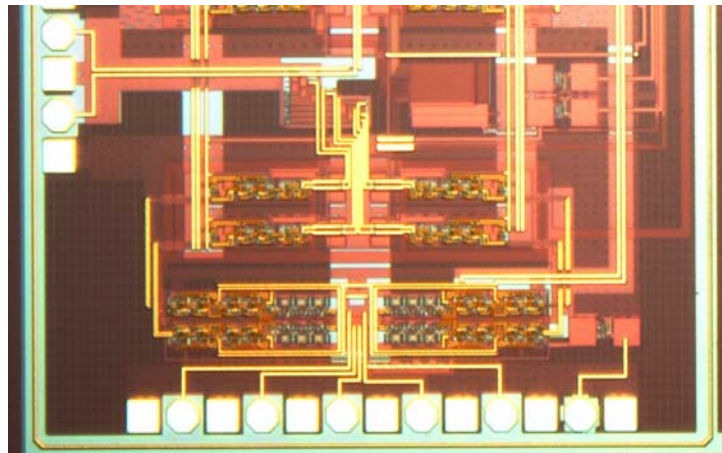
Fig. 5. 27. Chip micrograph of the 4-bit 30GHz DAC
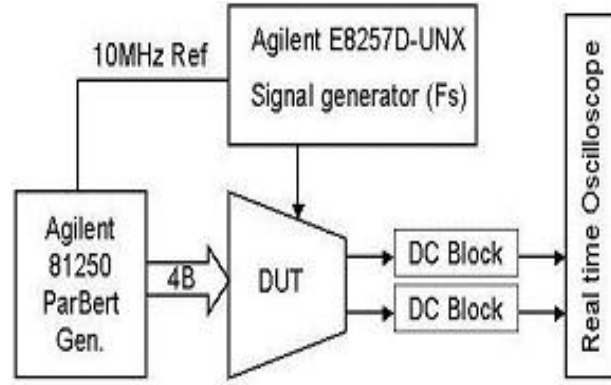
Fig. 5. 28. Measurement setup for the 4-bit 30GHz DAC

The 4-bit binary weighted DAC was tested on-wafer with a 40GHz probe station. For critical inputs and outputs 40GHz coaxial cables were used. The test setup is presented in Fig. 5.28. A low phase noise sinusoidal signal from an Agilent E8257D with option UNX was used as input clock. Since the output load of the DAC is matched with the external 50Ω load, it was possible to connect the outputs directly to the Tektronix 6154 oscilloscope through DC blockers. The input bit pattern is generated by an Agilent 81250 parallel bit-error rate tester, which was configured as a bit sequence generator. Unfortunately the module used for the characterization can only generate bit rate ≤ 3.35GHz. Thus the DAC could not be tested at the highest input data rate. By measuring the static and dynamic characteristics at lower data rate the parameters have been extrapolated for the higher data rate.
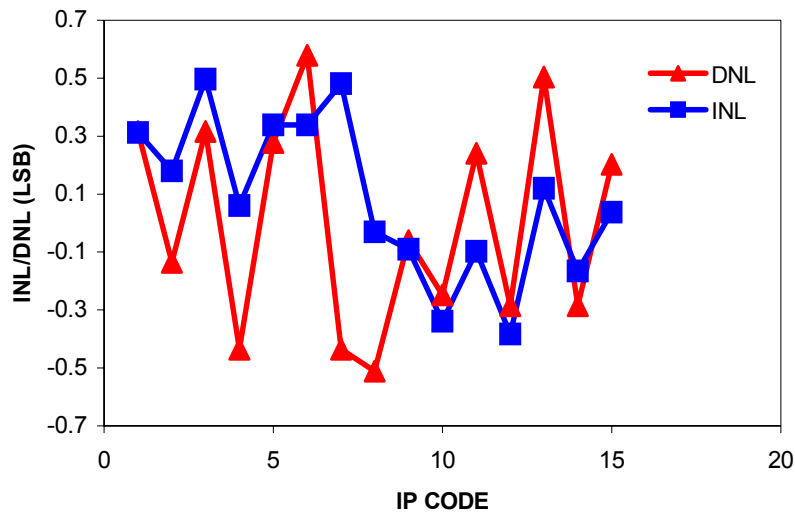
Fig. 5. 29. INL/DNL plot of 4-bit 30GHz DAC

In Fig. 5.29, the measured INL and DNL of the 4-bit DAC is plotted. It achieved INL and DNL of 0.49LSB and 0.57LSB respectively. Fig. 5.30a and 5.30b represents reconstructed

112

DAC output for different input bit patterns. Fig 5.30a shows the one of the differential output of the DAC for an input pattern corresponding to a sinusoidal function. With a data rate of 2.8GHz probe was observed. A full-swing step response of the DAC is presented in Fig. 5.30b with the input data rate of 500MHz and clock rate of 15GHz. Due to the lower cutoff frequency of the DC blocker (1GHz) the flat tops have some non-zero slope. . For the rise time measurement a reconstructed ramp signal is used. In Fig. 5.31a such a reconstructed ramp signal for the clock rate of 22GHz and 500MHz of input data rate is shown. The zoomed

(a)   (b)
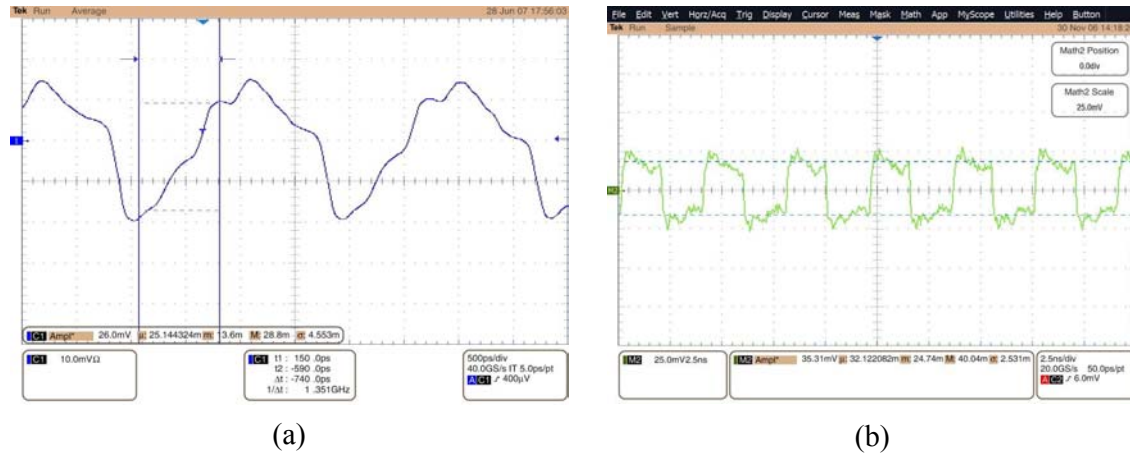
Fig. 5. 30. (a) Sinusoidal reconstruction for F_c=30GHz, I/P data rate=2.8GHz
(b) Step reconstruction for F_c=30GHz, I/P data rate=0.5GHz

portion of the full-scale transition is presented in Fig. 5.31b. From the rise time measurement (Fig. 5.31b) the output bandwidth of the DAC is calculated to be 3.85 GHz. Table 5.2 presents a summary of the measurement results.

(a)   (b)
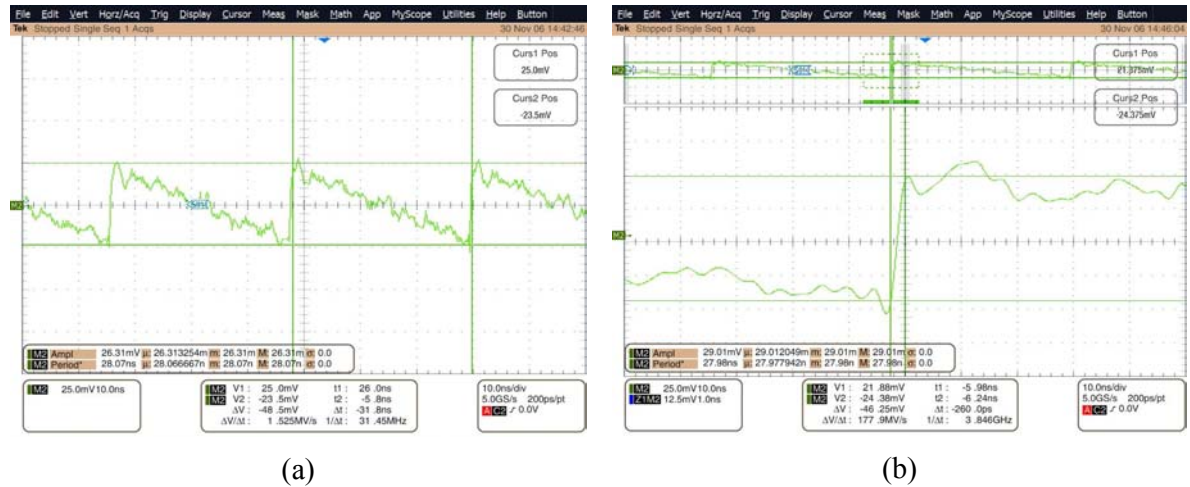
Fig. 5. 31. (a) Ramp reconstruction, (b) Rise time measurement
for Fc=22GHz, Data rate=0.5GHz

113

Table 5.2. Summary of measurement results

Process	IHP's 0.25µm SiGe BiCMOS SG25H1
Resolution	4 bit
Conversion rate	30 GHz
Output bandwidth	3.85 GHz
INL / DNL	0.49 / 0.57 LSB
Supply voltage	3.5 V
Power dissipation	455 mW
Die area with pads	1.87 mm^2

A common figure of merit (FOM) for a DAC relating sampling rate, power and resolution is expressed as,

$$FOM = \frac{Power}{2^N \cdot SamplingRate}$$

(5.1)

where, N is the resolution of the DAC. For the 4-bit DAC the FOM is 0.95pJ.

Table 5.3 shows a brief performance comparison among recently published high speed DACs in SiGe technology. The best FOM is found in [45], where special CML structures were used to reduce the power and the output bandwidth is very low (<1GHz). In [74] 0.13µm technology was used with reduced supply voltage and it shows the best performance in terms of sampling rate and resolution whereas the maximum sampling speed of 40GHz is achieved in [8]. This work the second highest sampling rate and has a comparable FOM at higher sampling rate in spite of higher supply voltage.

Table 5.3. Comparison with published Si/SiGe high speed DACs

Ref. No.	Fs [GHz]	Resolution [Bits]	Supply [V]	P_{diss} [mW.]	FOM (pJ)	Process / f_T (GHz)
[45]	20	6	1.8	360.0	0.28	0.18µm SiGe
[74]	22	6	-	1014	0.72	0.13µm SiGe/150
[74]	40	3	-	660.0	2.75	0.12µm Si/210
This work	30	4	3.5	455.0	0.95	0.25µm SiGe/190

5. 5. Conclusions

In this chapter design of two multi-GHz DAC has been presented. The 8-bit DAC has been implemented with a modified segmented current steering architecture. 50% segmentation is used to optimize the area resolution and the critical clock path length. The 4-bit LSB sub DAC is implemented with R-2R ladder network and the MSB sub-DAC has conventional unary weighted current steering architecture. In the unary weighed DAC design the thermometer decoder design comes as a bottleneck in terms of complexity, speed and power. A new architecture for the thermometer decoder has been proposed based on NOR/OR DFFs and the HBT ROM. In simulation the DAC shows 7.83ENOB for 9GHz of input sinusoidal with 20GHz of input clock.

A modified binary weighted current steering DAC is presented which can be used as a standalone DAC as well as a sub-DAC for a higher resolution segmented DAC. Unlike conventional binary weighted DACs, the weighting function is implemented in the load resistor instead of the current sources. The DAC achieves 0.49/ 0.57 LSB INL and DNL respectively with 3.85 GHz of output bandwidth. The DAC is found functional up to 30GHz of sampling rate. This is the second fastest DAC in SiGe technology according to the author knowledge. The DAC shows a FOM of 0.95pJ, which is comparable with the state-of-the-art SiGe high-speed DACs in spite of high power supply.

Chapter 6
Conclusions

_____

6. 1. Summary

In the last few decades the communication bandwidth has evolved with an enormous speed and the requirement of high-speed data converters is directly dictated by that. In RF systems, the analog-digital interface is pushed towards the antenna. Because the complex signal processing can be handled more efficiently in the digital domain. On the other hand it makes the design of these high-speed data converters more and more difficult.

 The scope of this current work involves the design of multi-GHz range data converter component designs. These components can be designed as standalone system and as be used to build up the multi-GHz data converter system.

In chapter 2 the different quantization processes are described. The static and dynamic errors associated with the quantization process are defined. The physical error sources which define the limit for the ADCs in terms of resolution and sampling rate are identified as the input referred thermal noise, the aperture uncertainty in the sampling process and the comparator ambiguity. The pros and cons of different ADC architectures suitable for multi-GHz sampling rate are analyzed. The flash architecture is found to be the fastest and most power hungry. An alternative to the flash ADC can be found in time interleaved architecture which is essentially the combination of number of parallel ADCs. The fastest sampling rate is achieved by using this architecture. But it comes with a large amount of digital post processing overhead which makes it unattractive for real time applications. In folding interpolating architecture a compromise can be found in terms of the speed, power and hardware complexity.

Design of different ADC components are presented in Chapter 3. In the context of multi-GHz ADC design, the front-end track and hold amplifier (THA) comes as the bottleneck for the full system. In this chapter different design techniques are presented to improve the performances of the THA so that the tough requirements of the quantizer block can be relaxed. Two different kinds of THAs are implemented and measured successfully. In both

the THAs, different techniques are used to enhance the input range up to 2Vpp differential at the sampling rate of 10GHz. To accomplish this requirement, the input buffers of the THAs are optimized. For the first time a cascode input buffer is used in the open loop THA design. This THA achieves 7.58 bits of accuracy at 10GS/s of sampling rate with 3 GHz of input bandwidth. Compared to the published high-speed THAs, the current work has better performance in terms of input range and bandwidth. At the same 2 Vpp swing, the improvement in ENOB is about three bits. According to the authors knowledge these THAs are the only published THAs which can work with 2Vpp input signal and achieve an accuracy of more than 6.5-bit at a sampling rate of 10GHz.

In the second implementation an emitter follower only THA circuit is presented. An adaptively V_{CE} adjusted npn pnp emitter follower is used as the input buffer to increases the input voltage swing. It achieves 6.2bits of accuracy at 10GHz of sampling rate with 1GHz of input bandwidth.

A new double sampled technique is proposed for the open loop THA architectures which can be instrumental to double the sampling speed of the THA with a little overhead of power dissipation compared to conventional open loop THAs. A novel double sampling switch is proposed which will make the sampling process insensitive to the clock skew, that appears as the bottleneck for the double sampling THAs and restricts the resolution.

As the basic building block of a quantizer an open loop comparator is designed, which can be used to build an 8-bit folding interpolating ADC. Measurement result shows that the comparator has 5.8-bit of resolution with the input bandwidth of 2GHz. Power dissipation of the core comparator is 70mW.

In the second part of the thesis the design of multi-GHz DAC has been presented. In chapter 4 different current steering DAC architectures have been presented. The static and error sources are analyzed. Different state-of-the-art techniques to enhance the performances of the current steering DAC are discussed. But those techniques are found to be not very efficient for the high-speed conversion range. A non-binary weighed DAC based current cell calibration technique is proposed which can be used for offline calibration of the current steering DAC with a very small area overhead.

In chapter 6 the design of two multi-GHz current steering DACs have been presented. A modified binary weighted current steering DAC is presented which can be used as a standalone DAC as well as a sub-DAC for a higher resolution segmented DAC. Unlike conventional binary weighted DACs, the weighting function is implemented in the load resistor instead of the current sources. The DAC achieves 0.49/ 0.57 LSB INL and DNL

respectively with 3.85 GHz of output bandwidth. The DAC is found functional up to 30GHz of sampling rate. This is the second fastest DAC in SiGe technology according to the author knowledge. The DAC shows a FOM of 0.95pJ, which is comparable with the state-of-the-art SiGe high-speed DACs in spite of high power supply.

The 8-bit segmented current steering DAC has already been designed, where the 4-bit 30GHz DAC is used as the LSB sub-DAC. The MSB sub-DAC is implemented with conventional unary weighted DAC architecture. In the context of high-speed DAC design the binary to thermometer decoder comes as the design bottleneck in terms of speed and power. In this unary sub-DAC design a novel thermometer decoder is proposed which is mainly based on an HBT ROM structure. In simulation the 8-bit DAC shows an accuracy of 7.83 effective number of bits (ENOB) with 9GHz of single tone input sinusoidal and a sampling rate of 20GHz.

6. 2. Future Works

In chapter 3 the design technique of clock skew insensitive double sampling THA has been presented. The simulation of the core THA is completed. The THA layout needs to be successfully completed and can be verified after the fabrication. The design of the main building blocks (e.g. the THA and comparator) for the ADC design are presented. These blocks can used to build up an 8-bit ADC system. A folding interpolating architecture would be most suitable for the ADC. Thus the folding interpolating amplifier has to be designed.

In chapter 5 the design of the 8-bit segmented current steering DAC has been presented. The 4-bit LSB sub-DAC is already designed and success fully measured. The fabrication of full 8-bit DAC has already been completed. But the measurement of the chip is not done yet which will be done very soon.

References

[1] R. V. D. Plassche; "Integrated Analog-to-Digital and Digital-to-Analog Converters"; Kluwer Academic Publishers; 1994.

[2] "IEEE Standard for Terminology and Test Methods for Analog-to-Digital Converters, Standard, Measurements"; IEEE Standard 1241-2000; Dec. 2000.

[3] B. E. Peetz; "Dynamic Testing of Waveform Recorders"; IEEE Trans. on Instrumentation and Measurement; vol. 32, no. 1, pp. 12–17; Jan. 1983.

[4] B. Razavi; "Principles of Data Conversion System Design"; IEEE Press, New York, 1995.

[5] R. H. Walden; "Analog-to-Digital Converter Survey and Analysis"; IEEE J. Selected Areas in Communications; vol. 17; pp. 539-550; Apr. 1999.

[6] B. Le, T. W. Rondeau, J. H. Reed, W. Bostian; "Analog-to Digital Converters"; IEEE Signal Processing Magazine; pp. 69-77; NOV. 2005.

[7] P. Schvan, D. Pollex, S. C. Wang, C. Flat, N. Ben-Hamida, "A 22GS/s 5b ADC in 130nm SiGe BiCMOS", Proc. IEEE ISSCC , pp. 572-573, 2006.

[8] W. Cheng et al., "A 3b 40GS/s ADC-DAC in 0.12μm SiGe", Proc. IEEE ISSCC, pp. 262-263, 2004.

[9] B. Razavi, B. A. Wooley; "Design Techniques for High-Speed, High-Resolution Comparators"; IEEE J. Solid-State Circuits, vol. 27, pp. 1916–1926, Dec. 1992.

[10] S. Tsukamoto et al., "A CMOS 6-b, 200MSample/s, 3-V Supply A/D Converter for PRML Read Channel LSI", IEEE J. Solid-State Circuits, vol. 31, pp. 1831–1836, Nov. 1996.

[11] K. Kattmann, J. Barrow; "A Technique for Reducing Differential Nonlinearity Errors in Flash A/D converters"; Dig. Tech. Papers International Solid-State Circuits Conference, pp. 170–171, Feb. 1991.

[12] K. Bult, A. Buchwald; "An Embedded 240-mW 10-b 50-MS/s CMOS ADC in 1mm^2"; IEEE J. Solid-State Circuits, vol. 32, pp. 1887–1895, Dec. 1997.

[13] Choi M., Abidi A. A., "A 6b 1.3GSample/s A/D Converter in 0.35μm CMOS," in Dig. Tech. Papers International Solid-State Circuits Conference, pp. 126–127, Feb. 2001.

[14] X. Jiang, Z. Wang, M. F. Chang.; "A 2GS/s 6b ADC in 0.18μm CMOS"; Dig. Tech. Papers International Solid-state Circuits Conference, pp. 322–323; 2003.

[15] U. K. Moon, G. C. Teams; "Digital Techniques for Improving The Accuracy of Data Converters "; IEEE Communication Magazine; pp. 957-965; OCT. 1999.

[16] S. H. Lewis, P. R. Gray; "A pipelined 5-Msample/s 9-bit analog-to-digital converter"; IEEE J. Solid-State Circuits, vol. 22, pp. 954–961, Dec. 1987.

[17] Vessal F., Salama C. A. T.; "An 8-Bit 2-Gsample/s Folding-Interpolating Analog-to-Digital Converter in SiGe Technology", IEEE Journal of Solid State Circuits, Vol. 39, pp. 238-241, 2004.

[18] T. Matsuura, T. Nara, T. Komatsu, E. Imaizumi, T. Matsutsuru, R. Horita, H. Katsu, S. Suzumura, K. Sato, "A 240-Mbps, 1-W CMOS EPRML Read-Channel LSI Chip Using an Interleaved Subranging pipeline A/D Converter," IEEE J. Solid-State Circuits, vol. 33, pp. 1840–1850, Nov. 1998.

[19] C. S. G. Conroy, D. W. Cline, P. R. Gray, "An 8-b 85-MS/s Parallel Pipeline A/D Converter in 1-μm CMOS," IEEE J. Solid-State Circuits, vol. 28, pp. 447–454, Apr. 1993.

[20] D. Fu, K. C. Dyer, S. H. Lewis, P. J. Hurst, "A Digital Background Calibration Technique for Time-Interleaved Analog-to-Digital Converters," IEEE J. Solid-State Circuits, vol. 33, pp. 1904–1911, Dec. 1998.

[21] K. Poulton et al., "A 20GS/s 8b ADC with a 1MB memory in 0.18μm CMOS," Proc. IEEE ISSCC, pp. 318-319, 2003.

[22] P. Vorenkamp and J. P. Verdassdonk, "Fully Bipolar 120-Msample/s 10-b Circuit", IEEE Journal of Solid State Circuits, Vol. 27, pp. 988-992, 1992.

[23] C. Fiocchi, U. Gatti and F. Maloberti, "Design Issues for High-Speed, High-Resolution Track-and-Hold in BiCMOS Technology", IEE Circuits Device and Systems, Vol. 147, pp. 100-106, 2000.

[24] Y. Borokhovych et al. "A Low-Power Track-and-Hold Amplifier in SiGe BiCMOS Technology", Proc. ESSCIRC, pp. 263-266, 2005.

[25] W. T. Colleran, A. A. Abidi, "A 10-b, 75-MHz Two-Stage Bipolar A/D Converter", IEEE Journal of Solid-State Circuits, Vol. 28, pp. 1187-1199, 1993.

[26] B. Razavi, "A 200-MHz 15-mW BiCMOS Sample-and-hold Amplifier with 3V Supply", IEEE Journal of Solid- State Circuits, Vol. 30, pp. 1326-1332, 1995.

[27] Y. Lu et al., "An 8-bit, 12Gsample/secSiGe Track-and-Hold Amplifier", Proc. BCTM, pp. 148-151, 2005.

[28] S. Halder, S. Osmany, H. Gustat, B. Heinemann, "A 10Gs/S 2V$_{pp}$ Emitter Follower Only Track and Hold Amplifier in SiGe BiCMOS Technology", Proc. of International Symposium on Circuit & Systems, 2006.

[29] S. Halder, H. Gustat, C. Scheytt, "An 8Bit 10Gs/S 2V$_{pp}$ Track and Hold Amplifier in SiGe BiCMOS Technology", ESSCIRC 2006.

[30] M. Waltari, K. Halonen, "Timing Skew Insensitive Switching for Double-Sampled Circuits," Proc. IEEE International Symposium on Circuits and Systems, vol. II, pp. 61–64, May 1999.

[31] V. D Plassche, " Differential Sampler Circuit", U. S. patent US005510736A, 1996.

[32] S. Halder, H. Gustat, "Open Loop Double-Sampling Track and Hold", German patent file no. 10 2007 031 130.5-55 Germany, 2007.

[33] G. Hoogzaad, "Double input Buffer for Track-And-Holde Amplifier", U. S. patent US20010007434, 2001.

[34] B. Heinemann et al., "Novel Collector Design for High-Speed SiGe:C HBTs", Proc. IEDM, pp. 775-778, 2002.

[35] B. Heinemann et al. "Complementary SiGe BiCMOS", Electrochemical Society Proceeding, vol. 2004-07, pp.25-31.

[36] B. Pregardier, U. Langmann and W. Hillery, "A 1.2-GS/s 8-b Silicon Bipolar Track&Hold IC", IEEE Journal of Solid- State Circuits, Vol. 31, pp. 1336-1339, 1996.

[37] Xiangtao Li et al., "A 5-bit, 18 GS/sec SiGe HBT track-and-hold amplifier", Proc. Compound Semiconductor Integrated Circuit Conf., pp.101-104, 2005.

[38] J. Lee et al. "A 50GS/s Distributed T/H Amplifier in 0.18μm SiGe BiCMOS", Proc. IEEE ISSCC, pp. 466-467, 2007.

[39] S. Shahramian A. C. Carusone, S. P. Voinigescu, "Design Methodology for a 40-GSamples/s Track and Hold Amplifier in 0.18- μm SiGe BiCMOS Technology"; IEEE J. Solid-State Circuits, vol. 41, pp. 2233–2240, 2006.

[40] W. M. L. Kuo, et al., "A 32 Gsample/sec SiGe HBT Comparator for Ultra-High-Speed Analog-to Digital Conversion", Proc. BCTM, 2005.

[41] Y. Borokhovych, H. Gustat, "A 20 GSample/s 40mW SiGe HBT Comparator for Ultra-High-Speed ADC", ECS Transactions, pp. 937-943, Oct, 2006.

[42] M. J. Flanagan, G. A. Zimmerman, " Spur-Reduced Digital Sinusoid Synthesis ", IEEE Transaction on Communication, Vol. 43, pp. 2254-2262, 1995.

[43] D. C. Larson, "High Speed Direct Digital Synthesis Techniques and Applications", Proc. GaAs IC Symposium, pp. 209-212, 1998.

[44] M. El Said, J. Sitch, M. Elmasry, "A 0.5 /spl mu/m SiGe pre-equalizer for 10 Gb/s single-mode fiber optic links", Proc. ISSCC, pp.224-225, 595, 2005.

[45] D. Baranauskas, D. Zelenin, " A 0.36W 6b upto 20GS/s DAC for UWB Wave Formation", Proc. ISSCC, pp. 580-581, 675, 2006.

[46] J. Ketola, et al, "Trensmitter Utilising Bandpass Delta-Sigma Modulator and Switch Mode Power Amplifier", Proc. ISCAS, pp. 633-636, 2004.

[47] B. Schafferer, R. Adams, "A 3V CMOS 400mW 14b 1.4GS/s DAC for Multi-Carrier Applications" Proc. ISSCC, pp.360-361, 532, 2004.

[48] S. Halder, H. Gustat, " A 30 GS/s 4-Bit Binary Weighted DAC in SiGe BiCMOS Technology ", Proc. BCTM, pp 46-49, 2007.

[49] C. H. Lin. K. Bult, " A 10-b, 500-Msample/S CMOS DAC in 0.6mm^2 ", IEEE Journal of Solid State Circuits, Vol. 33, pp. 1948-1958, Dec, 1998.

[50] A. Hastings, "The Art of Analog Layout", 2nd edition, Pearson International Edition.

[51] K. Lakshimikumar, et al., "Characterization and Modeling of Mismatch in MOS Transistor for Precision Analog Design", IEEE Journal of Solid State Circuits, Vol. 21, pp. 1057-1066, Dec, 1986.

[52] K. Lakshimikumar, et al., "A Comment on: Characterization and Modeling of Mismatch in MOS Transistor for Precision Analog Design", IEEE Journal of Solid State Circuits, Vol. 23, pp. 296, Feb, 1988.

[53] M. J.M. Pelgrom, et al., "Matching Properties of MOS Transistors", IEEE Journal of Solid State Circuits, Vol. 24, pp. 1433-1440, Oct, 1989.

[54] C. Conroy, W. Lane and M. Moran, "A Comment. On 'Characterization and Modeling of Mismatch in MOS Transistors for Precision Analog Design'", IEEE Journal of Solid State Circuits, Vol. 23, pp. 294-296, Feb, 1988.

[55] J. Bastos, et al., '' A 12 bit Intrinsic Accuracy High Speed CMOS DAC" IEEE Journal of Solid State Circuits, Vol. 33, pp. 1959-1969, Dec, 1998.

[56] J. J. Wikner, N. Tan, "Modeling of CMOS Digital-to-Analog Converter for Telecommunication", IEEE Transactions on Circuit and Systems-II, Vol. 46, pp. 489-499, May, 1999.

[57] T. Chen, G. G. E. Gielen, " The Analysis and Improvement of a Current Steering DACs Dynamic SFDR-1: The Cell-Dependent Delay Difference", IEEE Transactions on Circuit and Systems-I, Vol. 53, pp. 3-15, Jan, 2006.

[58] M. Clara, A. Wiesbauer, W. Klatzer, " Nonlinear Distortion in Current-Steering D/A Converters Due to Asymmetrical Switching Errors ", Proc. ISCAS, pp. 285-288, 2004

[59] A. V. D. Bosch, et al., "A 10-bit 1-Gsample/s Nyquist Current-Steering CMOS D/A Converter", IEEE Journal of Solid State Circuits, Vol. 36, pp. 315-324, Dec, 2001.

[60] G. A. M. Van der Plas, et al. " A 14-bit Intrinsic Accurecy Q^2 Random Walk CMOS DAC", IEEE Journal of Solid State Circuits, Vol. 34, pp. 1708-1718, Dec, 1999.

[61] J. Deveugele, et al., " A Gradient-error and Edge-Effect Tolerant Switching Scheme for a High-Accuracy DAC", IEEE Transactions on Circuit and Systems-I, Vol. 51, pp. 191-195, Jan, 2004.

[62] R. Van De Plassche, " A Monolithic 14-bit D/A Converter", IEEE Journal of Solid State Circuits, Vol. 14, pp. 552-556, Jun, 1979.

[63] L. R. Carley, " A Noise-Shaping Coder Topology for 15+ Bit Converters", IEEE Journal of Solid State Circuits, Vol. 24, pp. 267-273, Apr, 1989.

[64] L. R. Carlay, J. Kenney, " A 16-bit 4th Order Noise Shaping D/A Converter ", Proc. CICC, May, 1988.

[65] H. Gustat, J. Borngraber, "NOR/OR register based ECL circuits for maximum data rate**,** Proc. BCTM, pp 90-93, 2005.

[66] M. Vesterbacka, et al., " Dynamic Element Matching in D/A Converters with Restricted Scrambling**,** Proc. ICECS, pp 899-902, 2000.

[67] J. Deveugele, M. S. J. Steyaert, " a 10-bit 250-MS/s Binary Weighted Current Steering DAC", IEEE Journal of Solid State Circuits, Vol. 41, pp. 320-329, Feb, 2006.

[68] H. Gustat, "Offset calibration of 10GHz Amplifier", Proc. ISTDM,73, 2004.

[69] M. Rodwell, "High Speed Integrated Circuit Technology, Towards 100GHz Logic", World Scientific, 2001, ISBN 981-02-4638-2.

[70] S. Halder, H. Gustat, " A 30 GS/s 4-Bit Binary Weighted DAC in SiGe BiCMOS Technology ", Proc. BCTM, pp 46-49, 2007.

[71] S. Halder, H. Gustat, C. Scheytt, A. Thiede, "20GS/s 8-BitCurrent Steering DAC in 0.25μm SiGe BiCMOS Technology", accepted in European Microwave Integrated Circuits Conference, Oct, 2008.

[72] S. Halder, H. Gustat, C. Scheytt, " A 20 GS/s 8-Bit Segmented Current Steering DAC in SiGe BiCMOS Technology ", Microwave Technology and Techniques Workshop, European Space & Technology Centre, Noordwijk, May, 2008.

[73] T. Chen, G. G. E. Gielen, " The Analysis and Improvement of a Current Steering DACs Dynamic SFDR-1: The Cell-Dependent Delay Difference", IEEE Transactions on Circuit and Systems-I, Vol. 53, pp. 3-15, Jan, 2006.

[74] B. Schvan, et al, " A 22 GS/s 6b DAC with Integrated Digital Ramp Generator", Proc. ISSCC, pp. 122-123, 588, 2005.

List of Publication(s) and Patent(s)

1. S. Halder, S. Osmany, H. Gustat, B. Heinemann, "A 10Gs/S 2V_{pp} Emitter Follower Only Track and Hold Amplifier in SiGe BiCMOS Technology", Proc. of International Symposium on Circuit & Systems, 2006.

2. S. Halder, H. Gustat, C. Scheytt, "An 8Bit 10Gs/S 2V_{pp} Track and Hold Amplifier in SiGe BiCMOS Technology", ESSCIRC 2006.

3. S. Halder, H. Gustat, "Open Loop Double-Sampling Track and Hold", German patent file no. 10 2007 031 130.5-55 Germany, 2007.

4. S. Halder, H. Gustat, " A 30 GS/s 4-Bit Binary Weighted DAC in SiGe BiCMOS Technology ", Proc. BCTM, pp 46-49, 2007.

5. S. Halder, H. Gustat, C. Scheytt, A. Thiede, "20GS/s 8-BitCurrent Steering DAC in 0.25µm SiGe BiCMOS Technology", accepted in European Microwave Integrated Circuits Conference, Oct, 2008.

6. S. Halder, H. Gustat, C. Scheytt, " A 20 GS/s 8-Bit Segmented Current Steering DAC in SiGe BiCMOS Technology ", Microwave Technology and Techniques Workshop, European Space & Technology Centre, Noordwijk, May, 2008.