Design of 10-bit Sample and Hold Amplifier

A Thesis
Submitted in partial fulfillment of the requirement for the award of degree
of

Master of Technology
in
VLSI Design and CAD

By
Mr. Harcharan Singh
Roll No.-6040407
Batch 2004-2006

Under the Guidance of

Mr. Kulbir Singh
Sr. Lecturer
T.I.E.T, Patiala

Electronics & Communication Engineering Department
Thapar Institute Of Engineering And Technology
(Deemed University) Patiala (Punjab)-147004 (INDIA)
June, 2006

CERTIFICATE

This is to certify that the thesis report entitled **"Design Of 10-bit Sample and Hold Amplifier"** submitted by **Mr. Harcharan Singh** in the partial fulfillment of the requirement for the award of the degree of **Master of Technology in VLSI Design & CAD** to **Thapar Institute of Engineering and Technology (Deemed University), Patiala,** is a record of candidate's own work carried out by him under our supervision and guidance. The matter embodied in this report has not been submitted in part or full to any other university or institute for the award of any degree.

Mr. Kulbir Singh
Sr.Lecturer
T.I.E.T, Patiala

Countersigned by:

Dr. R.S Kaler
Professor and Head,
ECE Department
T.I.E.T, Patiala.

Dr. T.P SINGH
Dean of Academic Affairs
T.I.E.T, Patiala.

Acknowledgement

I express my sincere gratitude to my guides**, Mr. Kulbir Singh**, Sr.Lecturer, Electronics and Communication Engineering Department at Thapar Institute of Engineering & Technology, for their valuable guidance, proper advice, painstaking and constant encouragement during the course or my work on this thesis.

I also feel very much obliged to **Dr. R.S. Kaler,** Professor & Head, Department of Electronics and Communication Engineering for his encouragement and inspiration for execution of the thesis work.

I am deeply indebted to my parents for their inspiration and ever encouraging moral support, which enabled me to pursue my studies.

I am also very thankful to the entire faculty and staff members of Electronics and Communication Engineering Department for their direct–indirect help and cooperation.

Harcharan Singh
(ROLL NO. 6040407)

ABSTRACT

The highest bandwidth signal that can be digitized by an analog-to-digital converter is often govern by the performance of a preceding sample-and-hold circuit. Open-loop sample and hold topologies generally provide the fastest implementation of the sampling function. However, the precision obtained with such configurations is typically much lower than can be achieved with alternative closed-loop architecture. In design where an MOS transistor is used as sampling switch, input-dependent charge associated with the fast turn-off of the switch is often the principal source of sampling error. This charge injection introduces a pedestal error ΔVs , in the hold mode that results in both nonlinearity and gain error. Moreover, the pedestal error is not well-controlled and is therefore difficult to compensate for using self-calibration techniques.

In this thesis, the fully differential sample and hold circuit has been designed and simulated on 1.25um CMOS technology in order to meet the given specifications. This design is free from the non-linearities such input-dependent charge etc as stated above that are present in open-loop or single ended configuration.
The circuit designed is capable of sampling the input in the range of 2-8MHz with the precision of 10-bits, having clock frequency of 20MHz.

TABLE OF CONTENTS

Chapter 1

Introduction

1.1 Perspective

Digital technology improvements of the last decade have gained entry into every facet of electrical information processing. Old ideas have been improved and new applications created. In particular, the traditionally analog dominated communications sector has undergone a digital signal processing (DSP) revolution. Modern DSP techniques allow for systems with increasing complexity and functionality that is not realistically possible with analog electronics. DSP technology growth in both wired and wireless systems is driven by anticipated consumer demand on data speed (bandwidth), system performance (bit error rate and encoding), data security (encryption) and increasing number of users. Additionally, wireless systems must often meet these challenges within an already allocated and crowded frequency spectrum.

The focus of this work is the analog front end of a high performance wireless communications receiver. In order to understand the particular demands placed on analog circuitry in this application, consider a general viewpoint first and then study a specific example.

Improvement of digital wireless communications has been fueled by modern digital modulation as well as data and channel coding techniques. The cost of this sophistication is increased digital computational accuracy and complexity. One result is the migration away from binary to M-ary modulation methods, which entail multiple levels of quantization in some combination of amplitude, phase, or frequency. The quantization levels must be produced and acquired accurately, necessitating a digital to analog converter (DAC) in the transmit path and and analog to digital converter (ADC) in the receive chain. Next, the requirement to handle larger numbers of users has led to

relatively narrower channels with large nearby interferers. Taking into account a large variation in received signal power levels due to the near-far and multipath propagation issues, the result is a large dynamic range requirement (maximum versus minimum signal levels).

Finally, allowing for wideband operation (simultaneous acquisition of several channels instead of just one) and other system complexities this begins to explain why high-speed high-accuracy data converters are needed. Interestingly and sometimes ironically, the implementation of high-performance DSP is enabled by high-performance analog technology.

Consider the linearity requirements of basestation electronics in GSM900 digital cellular telephony as a case study to further illustrate these points. First recall that the signal to noise ratio (SNR) of an ideal N-bit ADC in its Nyquist band ($\frac{1}{2} f_s$) is given by

$$SNR(dB) = 6.02N + 1.76 \tag{1.1}$$

The quantization noise spectral density is

$$N_Q\left(\frac{-dBFS}{Hz}\right) = SNR + 10\log\left(\frac{1}{2} f_s\right) \tag{1.2}$$

where f_s is the sampling frequency. To meet the GSM900 specification NQ must be -156dBFS or better, yielding various different possibilities for the number of bits N and sampling frequency f_s. In general, this places us in the 13-14 bits and 40-80 MSPS range. Further considerations show that, assuming quantization noise is the dominant system noise source, SNR of about 80dB (equivalent to 13-bits) is required. Examining next the

effect of a blocker on a 200kHz signal channel with the blocker 3MHz removed, we find that the maximum tolerable spur level is about -100dBFS. The analog front end must thus be large-signal linear to more than 16-bits. Note that to test for this performance criterion the analog circuit must be driven by a near full scale signal; otherwise all input levels will not be tested and potential spur sources not detected.

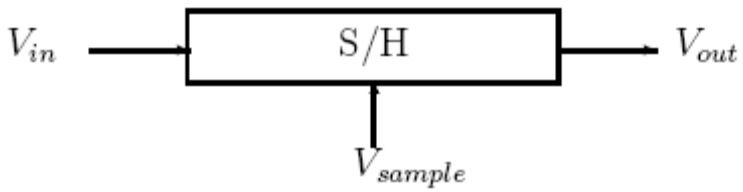


Figure1.1: S/H Block Diagram

The GSM example shows that on the receive side the analog front end (Sample and Hold and ADC) is a very critical block since it is responsible for sampling and quantizing the signal. High-speed, high-precision ADC's are the difference to allow a wideband implementation of the receive chain compared to several narrowband channels with substantially more hardware, cost, and power. To facilitate the high-speed quantization process, a high-performance Sample and Hold(S/H) circuit is a prerequisite.

Fundamentally there is a tradeoff between speed, power, resolution, linearity, noise and die area. Improving one specification means a sacrifice somewhere else.

1.2 The S/H Function

Conceptually, the premise behind the S/H circuit is simple. A time varying signal is fed to the S/H as well as a control signal as Figure 1.2 illustrates. While the control signal is active, the S/H tracks the input. When the control signal goes low, the last value of the input is held until the control signal is high again, starting the next track phase.

A distinction exists between S/H and Track and Hold (T/H) circuits. In most cases

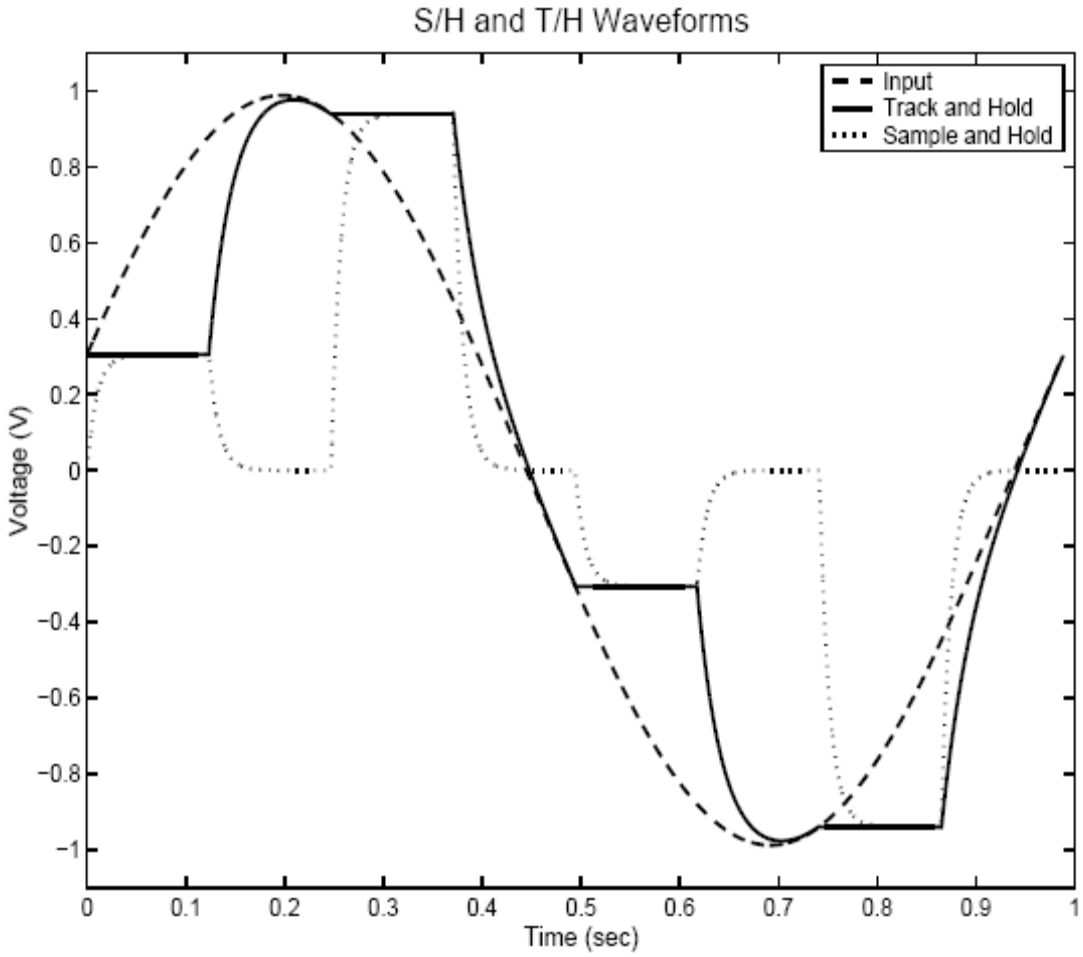


Figure1.2: The distinction between a S/H and a T/H

the terms are synonymous. However, S/H circuits do not necessarily have a phase of the output that follows, or tracks, the input. Instead, the output is reset to some predetermined value as shown in Figure 1.2. This may be a significant point for an under-sampled or a Nyquist rate sampled signal because in the worst case a T/H has to travel twice the distance compared to a S/H reset to the middle of the output range.

1.3 Design Goals

This work is totally intended for the realization of most widely used Sample and Hold circuit whose model was first represented by Razavi [2].

The 10-bit accuracy has been able to achieve with sampling frequency of 200 MHz. The fully differential sample and hold circuit has been realized which is free from most of the general non-linearities namely input dependent sampling time, charge injection etc that could have been present if single ended configuration would have been used.

Many other sample and hold circuit configurations have been proposed but the work presented here is the most general and widely used configuration.

Table 1.1 shows small view of the specification used in this work.

Table1.1 Design goals

Accuracy	10-bits
Power	Minimum(<2mW)
Input voltage Range	0.5-3Volts
Sampling Frequency	200Mhz

1.4 Literature Survey

To add some perspective, Tables 1.2 and 1.3 list state of the art S/Hs with ADCs performance for both commercial and published designs. Sampling rates range from 50 to 125 MSPS with ADCs of 12 to 14 bit resolutions.

Table1.2: State of the art Commercial S/H with ADC performance

Part	Vcc	Resolution	fs	Fin	SNR	Technology
AD6644[24]	5V	14-bit	65MSPS	30.5Mhz	73dB	Bipolar
AD6645[25]	5V	14-bit	80MSPS	30.5Mhz	74dB	Bipolar
AD6645[26]	5V	14-bit	105MSPS	70.0Mhz	72dB	Biploar
AD9432	5V	12-bit	105MSPS	50.3Mhz	67dB	BiCMOS
AD9433	3.3V	12-bit	125MSPS	70Mhz	64dB	BiCMOS

Table1.3 State of the art Published S/H with ADC performance

Authors	Vcc	Resolution	fs	fin	SNR	Technology
C. Moreland [7]	5V	14-bit	100MSPS	2.38MHz	75dB	Bipolar
L. Singer [13]	5V	12-bit	65MSPS	35MHz	69dB	0.5um CMOS
H. Pan [11]	3.3V	12-bit	50MSPS	25MHz	64dB	0.5um CMOS
D. Kelly [9]	3V	14-bit	75MSPS	37.5Mhz	72.8dB	0.35um CMOS

1.5 Organization of Thesis

In Second chapter presented the basic sampling fundamentals using mathematical equations have been presented. This chapter is only used to give the basic knowledge or an overview of sampling.

The third chapter talks about the CMOS sampling which includes all the non-linearities presented in the CMOS switch and sample and hold circuit. It also presents the solution to overcome some of the non-linearities.

Different sample and hold configuration has been discussed in Fourth chapter which also gives a brief introduction to the configuration used in this work and their trade-offs.

Fifth chapter discusses the whole design procedure and all the specifications used for the intended configuration and the results obtained are discussed.

Sixth Chapter discusses the conclusion and the future scope of sample and hold circuit.

CHAPTER 2

Mathematical Representation of Sampling

In order to thoroughly design a sampling circuit, an in depth understanding of the sampling process is requisite. This includes insight into mathematical models for both the time and frequency domains. An intuitive relationship between the the time and frequency domains is made possible by the notion of impulse sampling [1] which leads to the sampling criteria under which it is possible to reconstruct the original signal from its samples, the well-known Nyquist sampling criterion. Although the Nyquist criterion is usually only applied to baseband signals, recently there has been attention to under-sampling signals at IF frequencies greater than the ADC sampling rate. Because of the high input frequency, further demands are placed on the S/H which must be able to operate under these conditions as well. This chapter explains both `true' Nyquist sampling and under-sampling as well as the most important effects on the S/H.

2.1 Time Domain Representation

Given a continuous time signal $x_a(t)$, the sampling operation will produce the discrete time sequence $x(n)$ such that

$$x(n) = x_a(nT) \tag{2.1}$$

where $n \varepsilon N$ and $T = \frac{1}{f_s}$ where f_s is the sampling rate. A S/H does not produce $x(n)$ because its output is a continuous time signal that may represented by

$$x_{S/H}(t) = \sum_{n=-\infty}^{n=\infty} x_a(nT)[u(nT) - u((n+1)T)] \tag{2.2}$$

where $u(T)$ is the unit step function. (2.2) represents an ideal S/H with zero acquisition time, no hold mode droop, hold step or other non-ideal characteristics. An

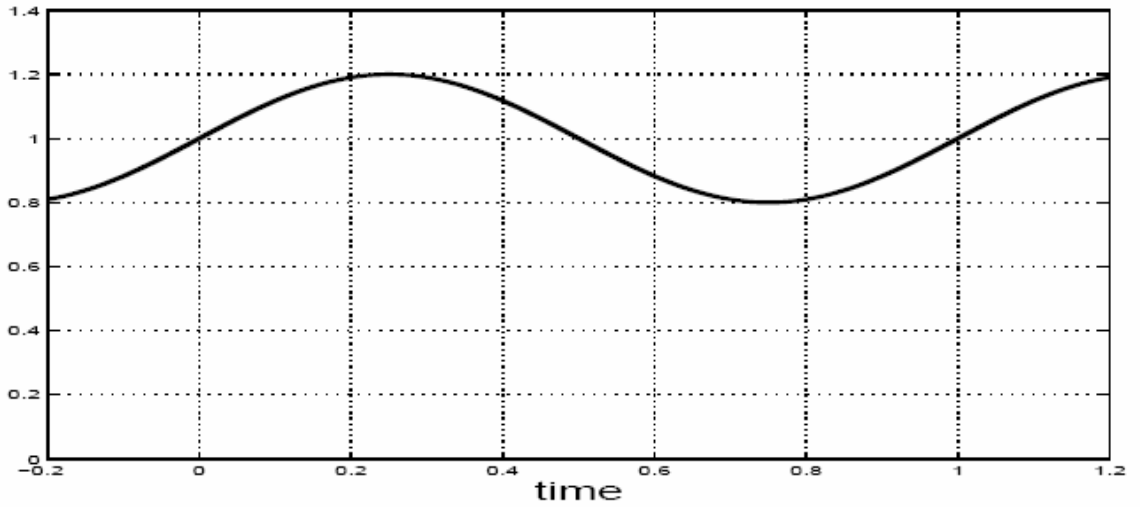


Figure 2.1: Analog signal to be sampled: xa(t)

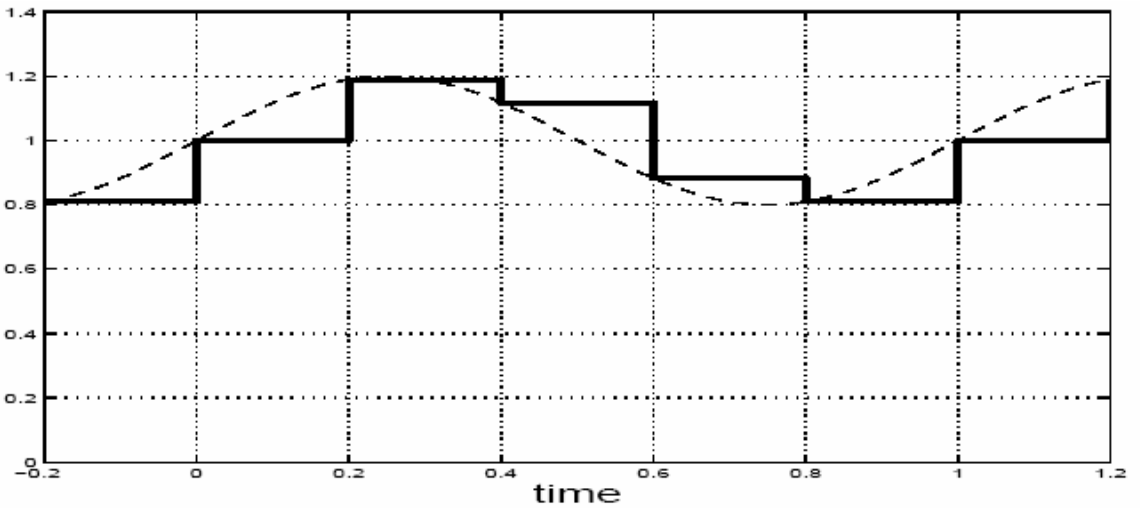


Figure 2.2: Ideal S/H version: $x_S = H(T)$

approximation to (2.1) is made by the ADC core that follows the S/H, with the difference being that $x(n)$ is now quantized also. Figure 2.1 through Figure 2.3 show $x_a(t), x_{S/H}(t)$ and $x(n)$. The dashed lines in Figure 2.2 and Figure 2.3 show the difference between

$x_a(t), x_{S/H}(t)$ and $x(n)$. It should be noted that in Figure 2.3, the envelope is shown as the dashed line for clarity.

2.2 Impulse Sampling

Impulse sampling is a fictitious operation with no real world equivalent. However, because of how it elegantly flows from a time domain representation to the frequency domain, a level of understanding of the effects of sampling on a signal not easily obtained from a rigorous derivation is facilitated. The major distinction between an impulse sampled.

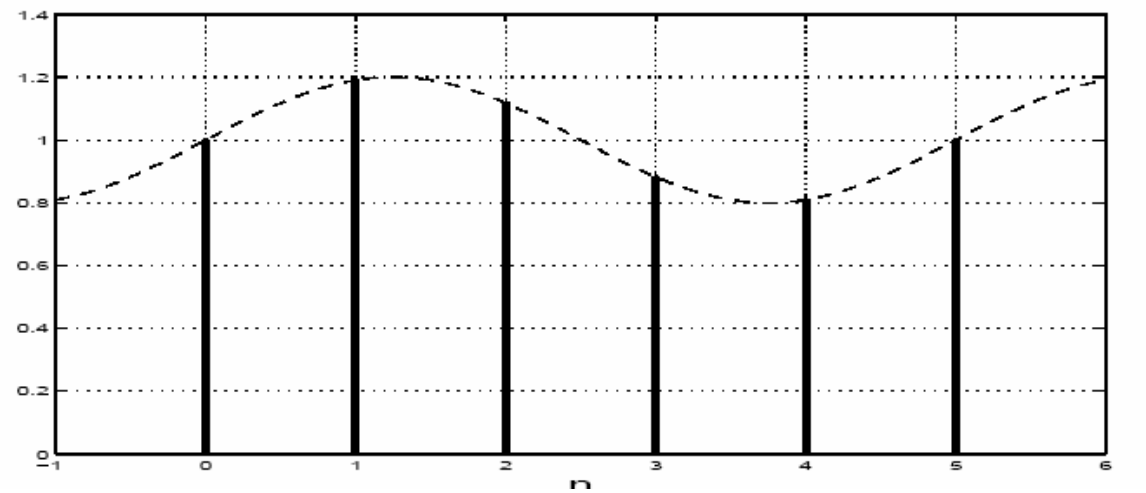


Figure 2.3: Discrete time sampled: x(n) = xa(nT)

signal and a discrete time sequence is that the former is defined for all time and the latter is a sequence evenly spaced at integer multiples with time information implied by the sampling rate.

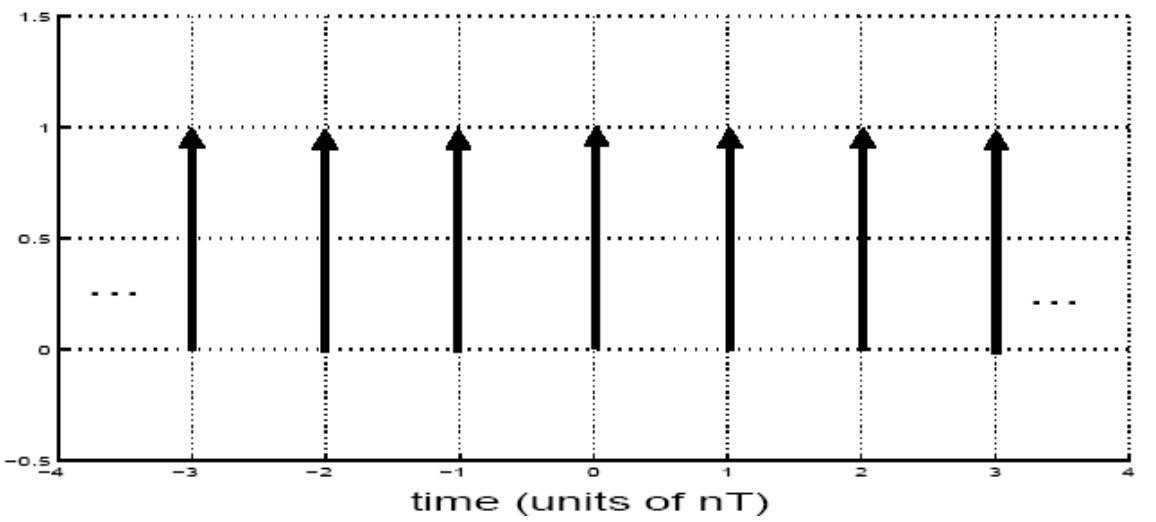


Figure 2.4: Impulse train sequence

Figure 2.4 shows an impulse train sequence that is used as the modulating signal in impulse sampling. Mathematically this may be expressed as

$$\delta_t(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT) \tag{2.3}$$

where T is the sampling period. The system in which this is used is shown in Figure 2.5.

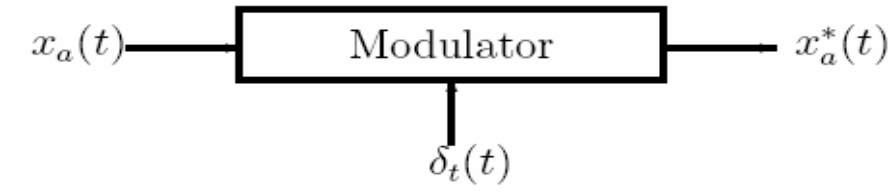


Figure 2.5: Impulse train modulator

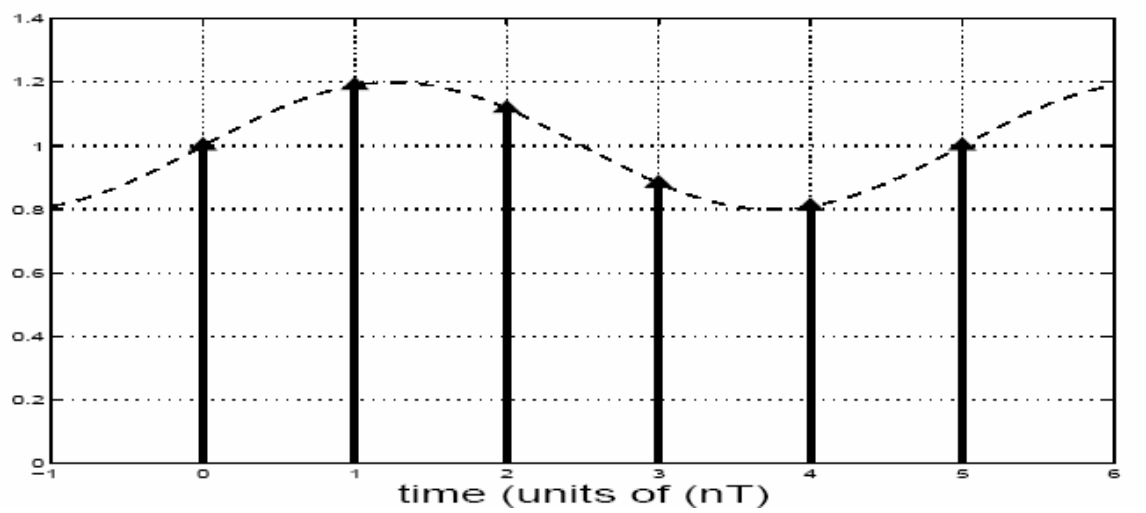


Figure 2.6: Impulse train modulated version of $x_a(t)$

$x^*_a(t)$ is related to $x_a(t)$ as follows:

$$x^*_a(t) = x_a(t)\delta(t)$$

$$= x_a(t)\sum_{n=-\infty}^{\infty}\delta(t-nT) \qquad (2.4)$$

(2.4) may be simplified to

$$x^*_a(t) = x_a(t)\sum_{n=-\infty}^{n=\infty}\delta(t-nT)x_a(nT) \qquad (2.5)$$

$x^*_a(t)$ produces impulses at spacings of nT in the continuous time domain with areas equal to $x_a(nT)$ respectively as seen in Figure 2.6. In Figure 2.6, $x_a(t)$ is superimposed with a dashed line.

2.3 Frequency Domain Representation

By taking the Fourier Transform, the frequency response is obtained. For $x_a(t)$ and $\delta_t(t)$, the Fourier transform pairs are $X_a(j\Omega)$ and $\delta_t(j\Omega)$ respectively.

Evaluating the Fourier Transform of $\delta_t(t)$ yields

$$\delta_t(j\Omega) = \frac{2\pi}{T}\sum_{k=-\infty}^{k=\infty}\delta(\Omega - k\Omega_s) \tag{2.6}$$

Ω_S represents the sampling frequency in radians per second. Since $x_a(t)$ and $\delta_t(t)$ are multiplied in the time domain, the Fourier Transform of their product, $x^*_a(t)$,is the convolution of their Fourier Transforms, $X_a(j\Omega)$ and $\delta_t(j\Omega)$, in the frequency domain,

$$X_a*(j\Omega) = X_a(j\Omega) * \delta_t(j\Omega) \tag{2.7}$$

Evaluating (2.7) produces

$$X_a*(j\Omega) = \frac{1}{T}\sum_{k=-\infty}^{\infty} X_a(j\Omega - k\Omega_s) \tag{2.8}$$

Using (2.8), relationships between the spectrum $X_a(j\Omega)$ with $X^*_a(j\Omega)$ can be deduced. The spectrum of the impulse modulated version of the input is the same as $X_a(j\Omega)$ with the addition of replicas at integer multiples of Ω_S scaled by $\frac{1}{T}$. Plotting $X^*_a(j\Omega)$ clearly illustrates the conditions under which a signal may be reconstructed.

$X_a(j\Omega)$ is shown in Figure 2.7 (a) having a maximum frequency of Ω_m radians per second. When $\Omega_m < \frac{1}{2}\Omega_s$, there is no overlap, as shown in Figure 2.7(b). As -m is increased, the replicas of $X_a(j\Omega)$ move closer to each other until $\Omega_m = \frac{1}{2}\Omega_s$ at which point the boundary for overlap is hit, shown in Figure 2.7 (c). Because there is no overlap, no aliasing has occurred and complete reconstruction is possible. However, when $\Omega_m > \frac{1}{2}\Omega_s$ Figure 2.7 (d) the replicas overlap and it is no longer possible to recover the original signal via an ideal low pass filter with a cutoff frequency of $\frac{1}{2}\Omega_s$.

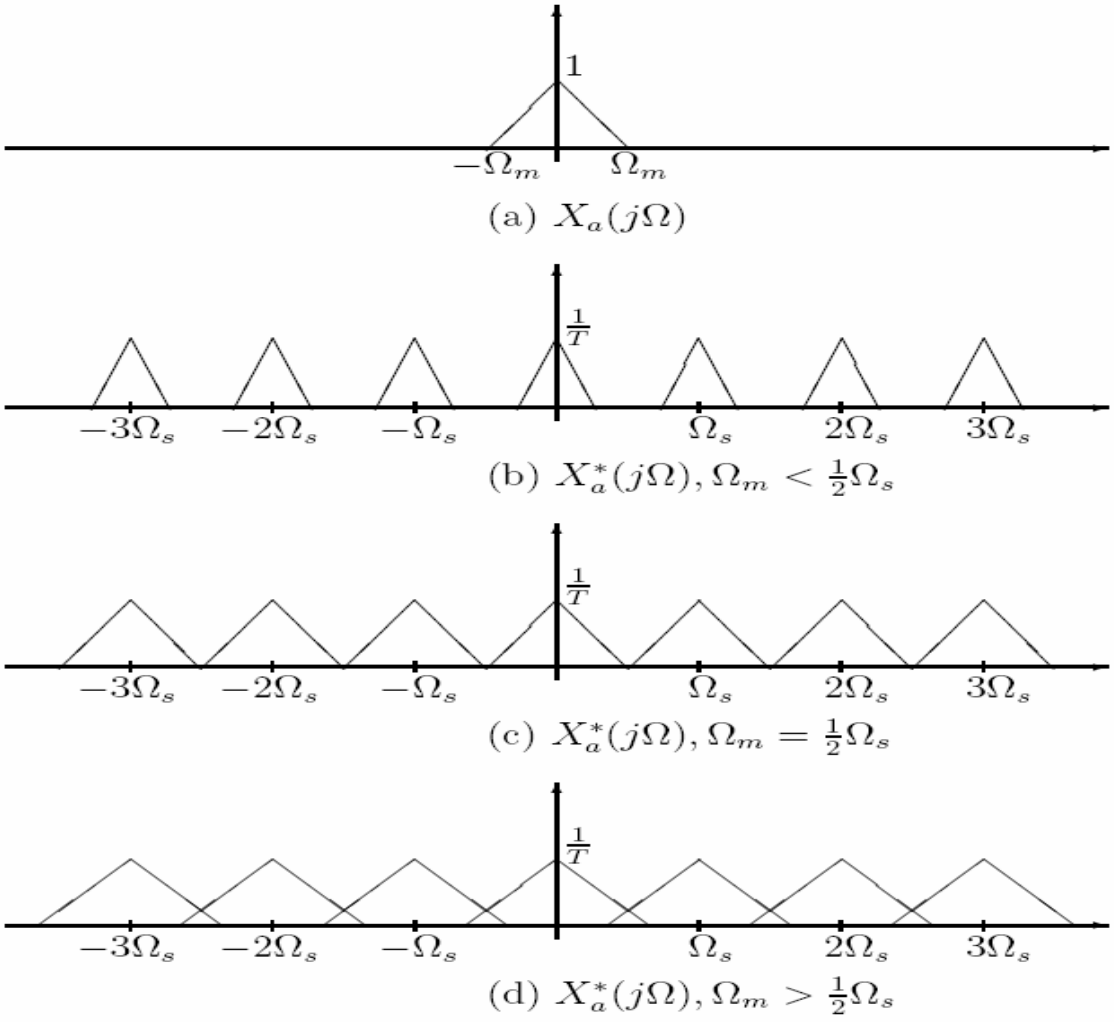


Figure 2.7: Baseband Spectrums

2.4 Under-sampling

Under special circumstances a signal may be under-sampled while avoiding aliasing. Two things are accomplished at once. First, the signal is sampled and therefore converted from continuous time to discrete time. At the same time, the signal is mixed down from an IF to DC. Simultaneously mixing the signal down to DC is significant because, on the

Table 2.1: Under-sample examples

Case	Center Frequency	Bandwidth
1	Ω_s	$\frac{1}{2}\Omega_s$
2	$\frac{3}{4}\Omega_s$	$\frac{1}{2}\Omega_s$
3	$\frac{3}{2}\Omega_s$	$\frac{1}{2}\Omega_s$
4	$\frac{5}{4}\Omega_s$	$\frac{1}{2}\Omega_s$

system level, the amount of hardware is reduced and a high quality mixer is not a trivial design. For this reason, being able to sample a signal while eliminating one or more IF down-conversions represent a benefit on the system level in many cases, especially in communication systems. Higher sampling rates are required for this to become a reality, shifting emphasis and therefore design difficulty to the S/H and ADC. This means a larger sampling bandwidth and clock jitter that is a fraction of the carrier (not clock) frequency. Furthermore, noise at the S/H input is folded into baseband (or new IF band), impacting SNR unless additional filtering before sampling is instituted. On the other hand (assuming open-loop sampling), the amplifier settling requirement is set by the clock rate, not the (higher) input frequency. Using the same simple graphical analysis that was used to attain insight into the effects of sampling the baseband, understanding of the conditions in which an under-sampled signal may be recovered without aliasing is possible. The Nyquist criterion still applies, therefore limiting the bandwidth of the signal to $\frac{1}{2}\Omega_s$.

Table 2.1 lists four examples that clarify this situation. Each case has a bandwidth of $\frac{1}{2}\Omega_S$ about a center frequency Ω_{centre}. The band is assumed to be symmetric with limits of $\Omega_{centre} \pm \frac{1}{2}\Omega_S$

Because aliasing occurs is some of the cases, the shape of the spectrum has been changed from a triangular to trapezoidal shape for clarity.

Case 1 has a center at Ω_S with the frequency response, $X_a(j\Omega)$, shown in Figure 2.7(a). When impulse sampled, Figure 2.7 (b), $X_a(j\Omega)$ is convolved with $\delta(j\Omega)$ producing scaled replicas of $X_a(j\Omega)$ at integer multiples of Ω_S. These replicas completely overlap each other making reconstruction impossible.

Given the symmetrically distributed bandwidth about -center, a shift away from Ω_S is needed to eliminate aliasing from Case 1. Figure 2.7 (c) and (d) demonstrate one such instance.

Case 2 has a center frequency of $\frac{3}{4}\Omega_S$ with the same bandwidth of $\frac{1}{2}\Omega_S$. Each replica of $X_a(j\Omega)$ is unaffected by adjacent replicas allowing reconstruction.

With results similar to Case 1, Case 3, Figure 2.8 (a) and (b), has no possibility for un-aliased reconstruction. The aliasing occurs due to overlapping spectra of replicas with an offset of $\pm 3\Omega_S$. In fact the overlap covers the entire bandwidth of $\frac{1}{2}\Omega_S$.

Shifting Ω_{centre} by $\frac{1}{4}\Omega_S$ to the right of that in Case 1, Case 4 demonstrates a second instance under which under-sampling is feasible. Like Case 2, replicas interleave just right so that adjacent sidebands do not overlap. $X^*_a(j\Omega)$ for Case 4, Figure 2.8 (d), at a glance appears to be identical to that of Case 2, Figure 2.8 (d). However, this is not the case. For the strip centered at DC, components in Case 3 are due to the $\pm\Omega_S$ replicas whereas in Case 4 the same components are a result of impulses at $\pm 2\Omega_S$.

A closed form expression of frequency bands that allow under-sampling without aliasing in terms of a center frequency -center about which a bandwidth of $\frac{1}{2}\Omega s$ is distributed is possible:

$$\Omega_{centre} = \pm\left(\frac{1}{4} + \frac{1}{2}n\right)\Omega s \qquad \forall n \in [0,1,2,3.....] \tag{2.9}$$

As the $\pm$ in (2.8) implies, the bands exist in both positive and negative frequency without which a real input sequence would be precluded. For the case where n = 0, there is no under-sampling since the signal is in the baseband and the Nyquist criterion is met. In fact n = 0 represents traditional sampling conditions.

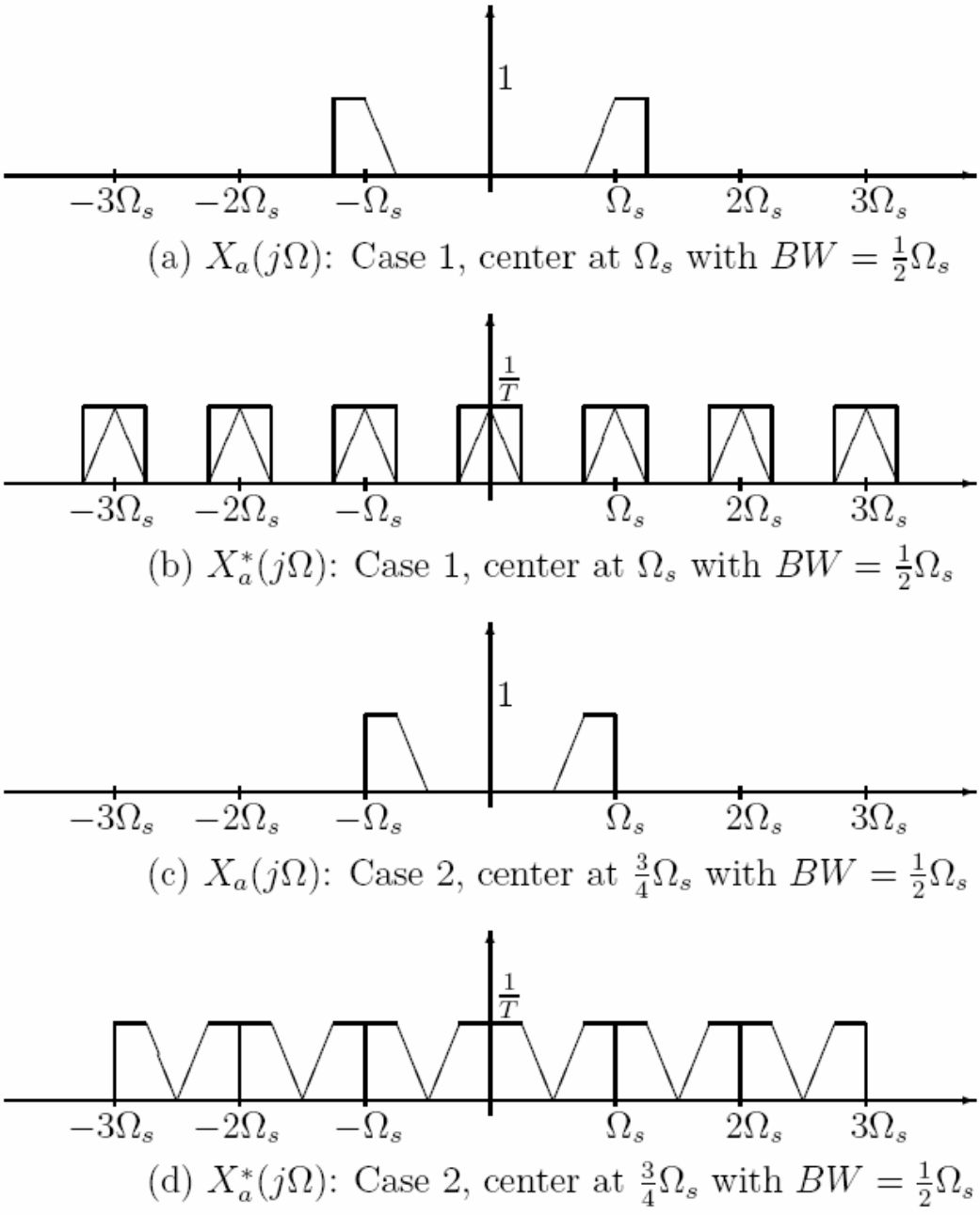


Figure 2.8: Effects of Under-sampling Part I

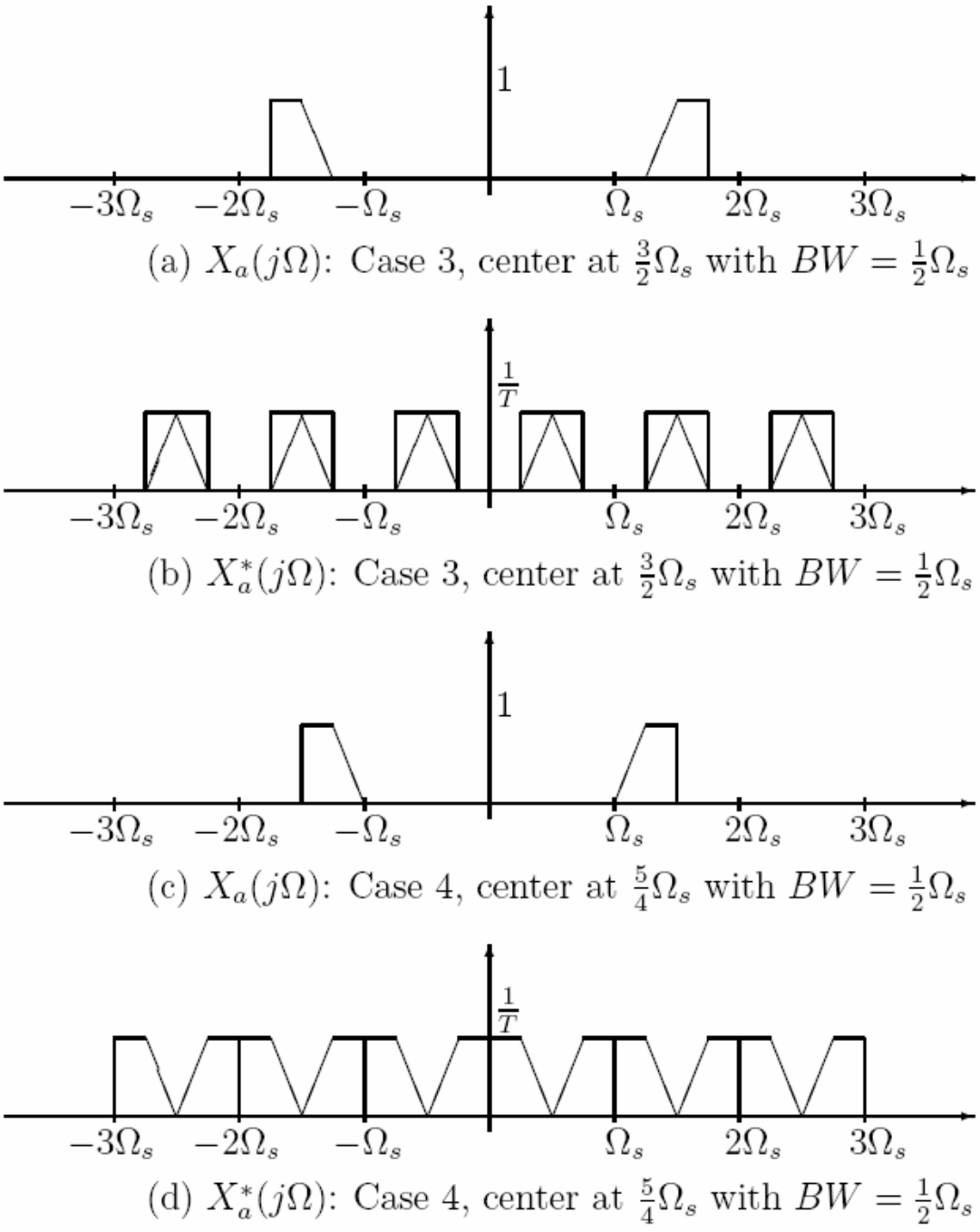


Figure 2.9: Effects of Under-sampling Part II

Chapter 3
Sampling Networks

For high-speed sampling, open-loop schemes are fundamentally better suited than closed loop schemes because closed-loop sampling techniques require quality amplifiers (Op-Amps) that are stable and fast. Hence, the maximum sampling rate is limited by the gain-bandwidth product of the amplifier. Conversely, open-loop sampling does not require an amplifier with stability restrictions, resulting in sampling networks with greatly reduced time constants compared to their closed-loop counterparts.

Sampling may be implemented in the voltage domain using either parallel or series techniques where the sampling capacitor is in parallel or series with the input voltage (Figure 3.1). Parallel sampling, Figure 3.1 (a), requires only a single switch (S_1) between the input and output nodes. When S_1 is shorted, V_{out} tracks V_{in} until the switch is opened, leaving the last value of V_{in} stored across CH in the form of charge. This charge is simply Q = CV. Series sampling, Figure 3.1 (b), is more complicated due to the need to employ two additional switches. In track mode, both S_2 and S_4 are shorted with S_3 open allowing node A to track the input with the output, node B, shorted to a convenient bias level (i.e. analog ground). In the transition to hold mode, S_4 is opened first, immediately followed by S_2 opening and S_3 shorting, connecting node A to ground. Consequently, a change at the output is produced equal to the last value of the input before the transition.

Some important distinctions exist between series and parallel sampling [2] [3]. Because of the orientation of the sampling capacitor, series sampling blocks the input common mode (DC) level from reaching the output, resulting in an additional degree of freedom since V_{out} and V_{in} may now operate at different bias points. Another advantage of series

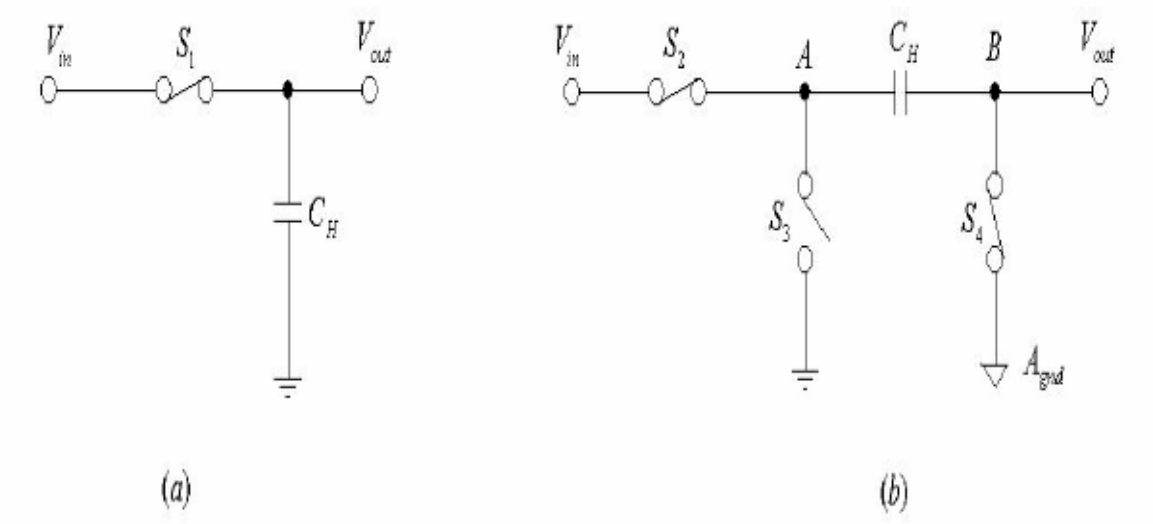


Figure 3.1: Sampling Methods: (a) Parallel, (b) Series

sampling is the elimination (in theory) of input dependent charge injection from the input switch (S_2) because the sample is taken when S_4 opens from the same voltage regardless of the input and therefore the charge injected is constant. Parallel sampling has the advantage of reduced hold mode feed-through since any parasitic capacitor that couples the input and output nodes across S1 are attenuated by CH. A disadvantage of series sampling is that a parasitic capacitor from node B to ground introduces a gain error into the system and therefore places a constraint on the minimum size that CH may take; if the parasitic capacitor is input dependent, nonlinearity results.

And in general the the sources of nonlinearity during sample (track) mode may be grouped as follows:

1) Channel Charge Injection
2) Clock Feedthrough
3) KT/C Noise
4) Non Linear Sampling Networks.

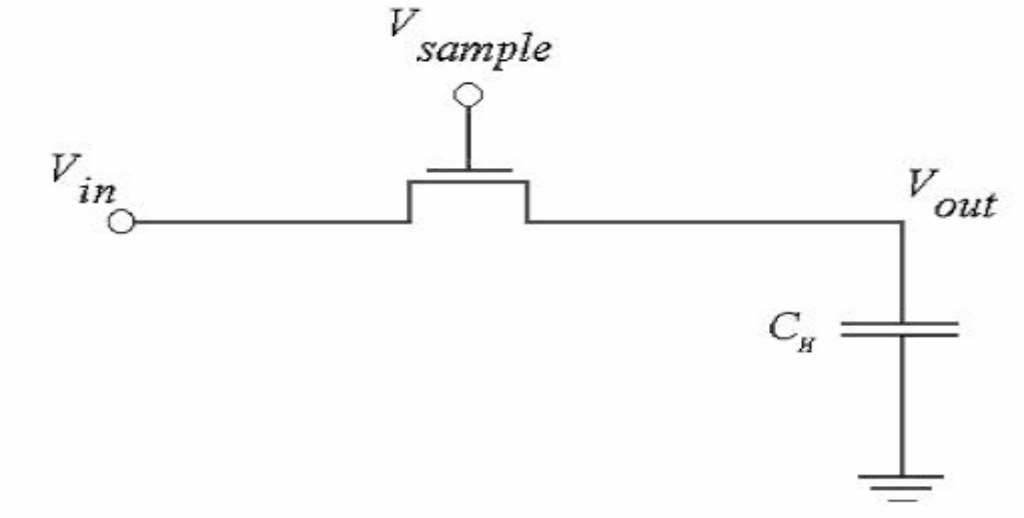


Figure 3.2: Simple CMOS S/H

3.1 CMOS Sampling

The basic element in any type of the sampling is the switch. Here in this section we will discuss the switch implementation of MOS, various capacitances associated with it and the non-linearties presents in the MOS switch

3.1.1 MOSFET as a Switch

In its simplest form, a CMOS S/H consists of just a MOS switch driven by a sampling clock signal and a capacitor (Figure 3.2); this represents parallel sampling and conveniently introduces switch non-idealities that may be extended to series or other more complex sampling methods. When Vsample is active, high for an NMOS and low for a PMOS, Vout follows Vin. At the end of the sampling period, Vsample abruptly changes to its off state, low for an NMOS and high for a PMOS, disconnecting Vin from Vout because, when off, a MOS device is very close to an open circuit. The last value of Vin is stored on the capacitor in the form of charge.

The performance of parallel sampling in CMOS is limited by the quality of the switch and requires insight into an MOS device operated in the linear region to understand why?

A commonly accepted model for a MOSFET in the triode region is shown in Figure 3.3 [4]. The substrate connection is assumed to be connected to ground. For an

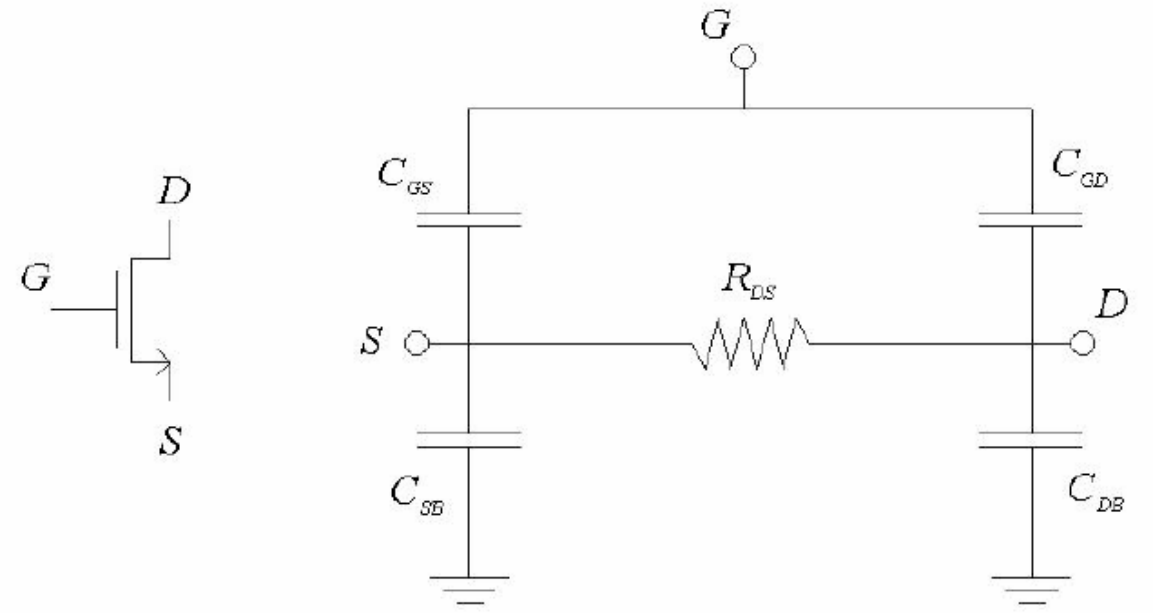


Figure 3.3: NMOS Small Signal Triode Region Model

NMOS device, the drain current is

$$I_{DS} = \mu_n C_{OX} \frac{W}{L}(V_{GS} - V_t - \frac{1}{2}V_{DS})V_{DS} \tag{3.1}$$

where, W is the width, L is the length, μ_n is the electron mobility, C_{OX} is the oxide capacitance per unit area, V_t is the threshold voltage, V_{GS} is the gate to source voltage, and V_{DS} is the drain to source voltage. Taking the partial derivative of (3.2) with respect to V_{DS} yields

$$\frac{\partial I_{DS}}{\partial V_{DS}} = \mu_n C_{OX} \frac{W}{L}(V_{GS} - V_t - V_{DS}) \tag{3.2}$$

Inverting Equation (3.2) results in the drain to source resistance

$$R_{DS} = \frac{L}{W\mu_n C_{OX}(V_{GS} - V_t - V_{DS})} \tag{3.3}$$

From (3.3), a MOSFET in the triode region acts as a voltage controlled resistance with the gate as the controlling node since V_{DS} is typically very small in track mode (except transiently). Because V_{GS} depends on the source voltage, R_{DS} is input dependent and introduces nonlinear distortion.

Small signal capacitors in the triode region model for a MOS device are due to both parallel plate and junction types. These are illustrated in Figure 3.4. C_{GD} and C_{GS} can be approximated as parallel plate capacitances that evenly divide the gate to channel capacitance between the source and drain. With L denoting the drawn length (i.e. including overlap) they can be approximated as

$$C_{gs} = C_{gd} = \frac{A_{ch}C_{OX}}{2} = \frac{WLC_{OX}}{2} \tag{3.4}$$

where, A_{CH} is the area of the channel. From both the drain and source to the bulk there are reversed biased pn junctions. Assuming an abrupt junction (which results in the square root in the denominator below), these capacitances may be may be estimated as [4]

$$C_{sb} = \frac{C_{sb-o}}{\sqrt{1 + \frac{V_{sb}}{\varphi_0}}} \tag{3.5}$$

where

$$C_{sb-0} = C_{jo}A_s + \frac{A_{ch}}{2}C_{jo-ch} + C_{j-sw0}P_s \tag{3.6}$$

and

$$C_{db} = \frac{C_{db-o}}{\sqrt{1 + \frac{V_{db}}{\varphi_0}}} \tag{3.7}$$

where

$$C_{db-0} = C_{jo}A_s + \frac{A_{ch}}{2}C_{jo-ch} + C_{j-sw0}P_d \tag{3.8}$$

C_{j0} is the zero bias depletion capacitance of a pn junction,

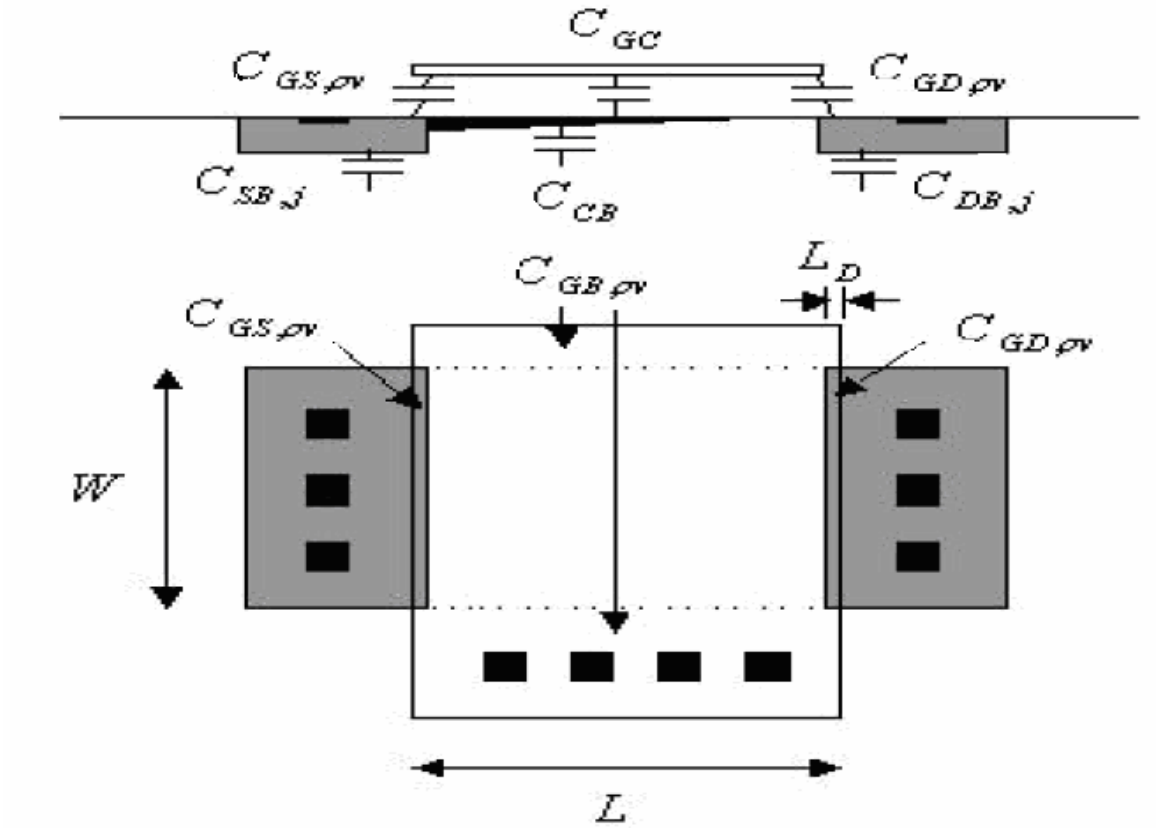


Figure 3.4: MOS Capacitances

and C_{j-sw0} is the zero bias depletion capacitance of its sidewall. For an abrupt junction

$$C_{j0} = \sqrt{\frac{qK_s\varepsilon_o N_A N_D}{2\varphi_0(N_A + N_D)}} \tag{3.9}$$

N_A and N_D are the acceptor and donor doping levels. ϕ_0 and q are the built-in voltage of an open circuit pn junction and the charge of an electron respectively. K_S and $\in_0$ are the relative permittivity of silicon and free space. From (3.6) and (3.8), both the drain to bulk and source to bulk capacitors are dependent on the amount of reverse bias voltage across each junction. Since V_{DS} is very small, both the drain as well as the source follow the input making both C_{SB} and C_{DB} input dependent and nonlinear, introducing a potential source of distortion for both track and hold modes.

In cutoff, a MOS device may be modelled by capacitances only Figure 3.5 since the source to drain resistance is now nearly infinite. Both C_{GS} and C_{GD} are considerably reduced from the triode region model because the channel is gone and they are now due

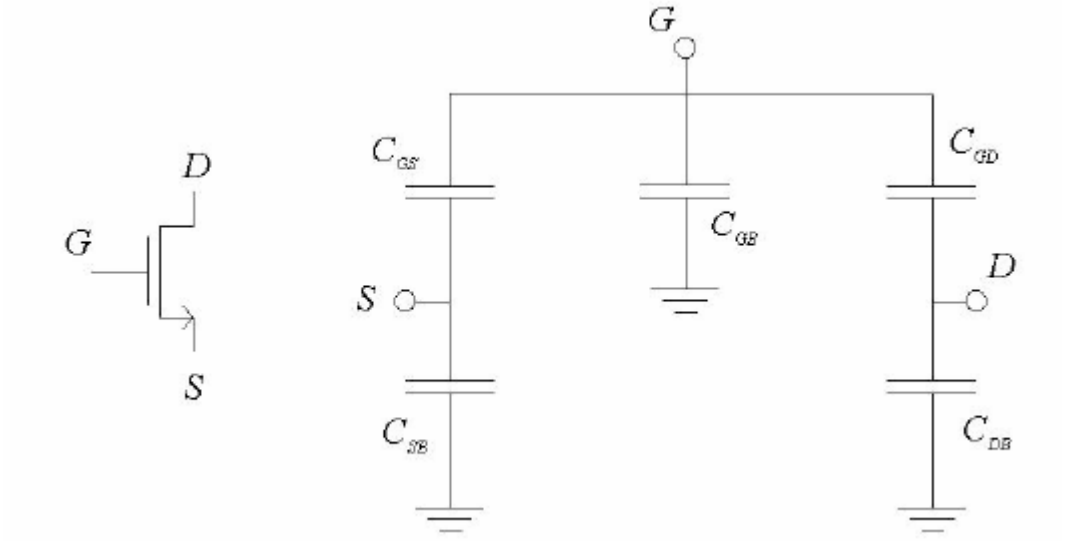


Figure 3.5: MOS Cutoff Model

to overlap and fringing. Combining both effects into one,

$$C_{GS} = C_{GD} = WL_{OV-eff}C_{OX} \tag{3.10}$$

A capacitor from the gate to the bulk (C_{GB}) is now present; due to the highly nonlinear nature of this capacitor, a worst case model is often used and occurs when the gate has been very negative for a considerable amount of time and the gate has accumulated:

$$C_{GS} = A_{CH}C_{OX} = WLC_{OX} \tag{3.11}$$

C_{SB} and C_{DB} are still given by Equations 3.6 and 3.8 except that the areas making up C_{SB-0} and C_{DB-0} lack the channel-bulk area, yielding smaller total capacitances:

$$C_{SB-0} = A_s C_{jo}$$

$$C_{DB-0} = A_s C_{jo} \tag{3.12}$$

3.1.2 Input Dependent Sampling Time

A MOS device is in cutoff when $V_{gs} \le V_t$. For the parallel MOS sampling circuit in Figure 3.2, both quantities are functions of the input. Consequently, the time at which the sample is taken (i.e. when the MOSFET just enters cutoff) depends upon the input signal and varies from one sampling period to the next.

Assuming a sinusoidal input of the form $V_{IN}(t) = A\sin(\omega t)$, this effect may be illustrated by examining two consecutive samples taken at t_1 and t_2. Ideally, $t_2 = t_1 + T$.
where T is the sampling period. In actuality, $t_2 = t_1 + T + \Delta T$ where ΔT is the deviation from the ideal sampling period.

Therefore

$$V_1 = A\sin(\omega t_1)$$

$$V_2 = A\sin(\omega t_2)$$

Then the deviation in V_2 from its ideal value is:

$$\begin{aligned}\Delta V_2 &= V_{2-Ideal} - V_{2-Actual} \\ &= A\sin(\omega(t_1+T)) - A\sin[\omega(t_1+T+\Delta t) \qquad (3.13)\end{aligned}$$

Applying trigonometric identities, assuming $\omega\Delta t << 1$ and simplifying (3.13) reduces to

$$\Delta V_2 \approx -A\omega\Delta T\cos(\omega(t_1+T)) \qquad (3.14)$$

(3.14) is at a maximum when $\cos(\omega(t_1+T)) = \pm 1$. Therefore, the maximum error is

$$\left|\Delta V_{2-\max}\right| = A\omega\Delta T = A\omega\frac{\Delta V_{IN}}{V_{CC}}T_F \qquad (3.15)$$

where V_{CC} is the supply voltage (the level for a high clock signal), ΔV_{IN} is the change in the input over one sampling period and T_F is the fall time of the clock signal. Because $\left|\Delta V_{2-\max}\right|$ is proportional to the input frequency, the error is greatest at the Nyquist frequency. Furthermore, it is also directly proportional to the fall time of the clock; the faster the fall time, the lower the error will be.

To further emphasize the importance of this effect, consider the following example: Assuming a sampling rate of 100 MSPS, the worst case will occur at the Nyquist frequency of $\omega = 2\pi 50$ Mrad/s with ΔV_{IN} = 2V. Also, assume a supply voltage of 5V. Then $\left|\Delta V_{2-\max}\right|$ = 12:6 mV. If the input amplitude is at full scale, then this represents an accuracy of only 7.02 bits. Hence, input dependent sampling time is a fundamental limit for parallel MOS sampling circuits.

3.1.3 Charge Injection

In the transition from track to hold mode, the MOS device goes from strong inversion in the linear region to cutoff. When this happens, the charge that is in the channel must leave since the condition that held it there, $V_{GS} > V_t$, no longer holds true [3]. This charge is simply

$$Q_{channel} = WLC_{OX}(V_{GS} - V_{tn} - V_{DS})$$

Charge distribution between the source and drain is usually assumed to occur according to the impedances seen at these terminals looking out. In direct proportion to these impedances, the charge will be distributed with the smaller impedance proportionally taking more of the charge. Since the amount of charge is dependent on V_{GS}, input dependent nonlinear charge injection results [4] [10] [12]. Further complicating the situation is that the threshold voltage V_t is also input dependent because of the back gate effect.

$$V_t = V_{t-0} + \gamma(\sqrt{V_{SB} + 2\phi_F} - \sqrt{2\varphi_F}) \tag{3.16}$$

where

$$\gamma = \frac{\sqrt{2qN_AN_DK_S\varepsilon_0}}{C_{OX}} \tag{3.17}$$

and

$$\Phi_F = \frac{KT}{q}\ln\left(\frac{N_A}{n_i}\right) \tag{3.18}$$

V_{t0} is the zero bias threshold voltage.

To reduce the amount of nonlinearity introduced, the length of the input switch must be minimized, the impedance of analog ground minimized also, pump overdrive and the reference voltage maximized. This leads to increased bias power consumption, the need for smaller geometry processes and larger supply voltages.

3.1.4 Clock feedthrough

Ideally, when the clock edges are switching, this will not affect the operation of SC circuit. However, due to some parasitic capacitance associated with the MOS switches, the voltage difference in clock switching may be coupled into the sampled input and create an error [4] [20].

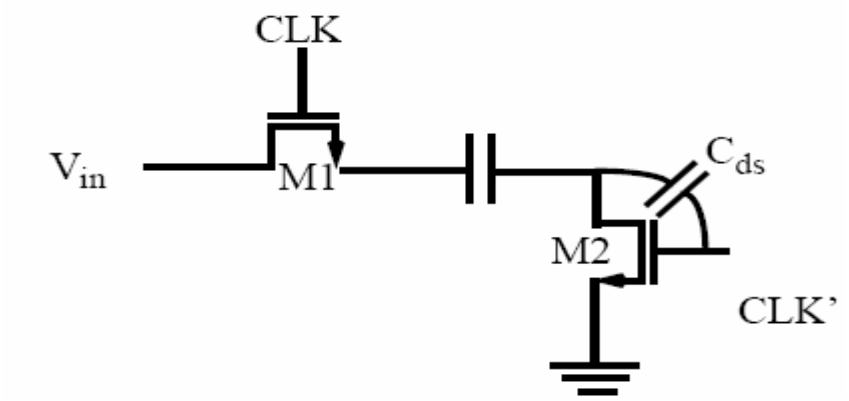


Figure 3.6: Illustrations for Clock Feedthrough

Figure 3.6 shows a typical sample-and-hold circuit employing the bottom plate sampling described in the previous section. C_{ds} is the overlapping capacitance between the gate and drain of the MOS switch. When the clock voltage on the gate switches between high and low, this voltage drop is coupled into the signal via the capacitor divider [3].

The clock feedthrough can be corrected to the first order by using a differential signal path. As long as the error is present on both signal inputs and the same magnitude, it can be cancelled by taking the input differentially. However, this technique, once again, depends on the absolute matching of transistors.

3.1.5 kT/*C* Noise

As mentioned earlier, during the "on" phase, the switch can be modeled as a resistor. The equivalent thermal noise of this resistor (R_{on}) has a one-sided, white-noise-like power spectral density of [8]

$$\frac{\overline{V^2}}{\Delta f} = 4kTR_{on} \tag{3.19}$$

where k is the Boltzmann's constant (1.38 × 10−23 JK−1), and T is the sampling clock period. As the switch is primarily used in a sampling network where a sampling capacitor (C_S) is charged by the input signal through the switch during the "on" phase, the thermal noise is processed by a first-order low-pass filter composed of R_{on} and C_S, whose transfer function is given by

$$H(j\omega) = \frac{1}{1 + j\omega R_{on} C_S} \tag{3.20}$$

Thus, the total noise power at the filter's output is computed by integrating the filtered (or bandlimited) noise power spectral density from dc to infinity, and we have

$$\begin{aligned}
\overline{V^2_{out}} &= \frac{1}{2\Pi}\int_0^{\infty} |H(j\omega)|^2 \overline{V^2_{in}} d\omega = \frac{1}{2\Pi}\int_0^{\infty} \frac{4kTR_{on}}{1 + (\omega R_{on} C_S)^2} \\
&= \frac{2kTR_{on}}{\Pi}\left(\frac{1}{R_{on}C_S}\right)\arctan(\omega R_{on} C_S)\,|_{\omega=0}^{\omega=\infty} \\
&= \frac{KT}{C_S}
\end{aligned} \tag{3.21}$$

hence the name *kT/C noise* (sometimes also called the *sampling noise*). Interestingly, the total noise power at the output is independent of the actual on-resistance value (R_{on}). This is because when the thermal noise power density is increased with the on resistance (*Ron*), the *noise bandwidth* of the filter, which is given by

$$f_n = \frac{\pi}{2} f_{-3dB} = \frac{\pi}{2} \frac{1}{2\pi R_{on} C_s} = \frac{1}{4 R_{on} C_s} \tag{3.22}$$

is decreased with the same magnitude. Specifically, the total noise power can be easily calculated by multiplying the spectral density given by Equation (3.19) by the noise bandwidth noted earlier, and we have

$$\overline{V^2} = \overline{V^2}_{\text{in}} f_n = 4kTR_{on} \frac{1}{4R_{on}C_s} = \frac{KT}{C_s} \tag{3.23}$$

which is the same as the result given by (3.21). Furthermore, for a given value of R_{on} , the total output noise power is decreased with the clock period *T*. That is, if the signal is oversampled by a factor of *M*, then the resultant kT/C noise power is also reduced by a factor of *M*. In a similar manner, oversampling (i.e., sampling at a rate much higher than the Nyquist rate) is useful for suppressing the white-noise-like quantization noise of an analog-to-digital data converter (ADC).

3.1.6 Nonlinear Sampling Network

In track mode, the input sampling network may be modelled as shown in Figure 3.7. This circuit contains only a resistor to model the switch on resistance and the sampling capacitor C. In order for an accurate sample to be taken, $V_{OUT}(t)$ must reach a steady state condition and precisely follow $V_{IN}(t)$. This circuit is described by the differential equation.

$$V_{IN}(t) = V_{OUT}(t) + \tau \frac{dV_{out}(t)}{dt} \tag{3.24}$$

where τ = RC. The complete solution of (3.24) given a sinusoidal input of the form

$$V_{IN}(t) = A\cos(\omega t)$$

consists of the sum of a homogeneous solution that satisfies

$$0 = V_{OUT}(t) + \tau \frac{dV_{out}(t)}{dt}$$

and a particular solution due to the input. Rewriting the input as $V_{IN}(t) = \Re\left\{A\,\omega e^{j\omega t}\right\}$ and assuming a solution of the form $V_{OUT-P}(t) = \Re\left\{k\omega e^{j\omega t}\right\}$ where k is a complex number, the particular solution must satisfy

$$\Re\left\{Ae^{j\omega t}\right\} = \Re\left\{ke^{j\omega t} + \tau j\omega k e^{j\omega t}\right\}$$

Thus,

$$k = \frac{A}{1 + j\omega\tau} = \frac{A}{\sqrt{1+\omega^2\tau^2}} e^{-j\theta} \tag{3.25}$$

where θ = arctan($\omega\tau$). Hence

$$V_{OUT-P}(t) = \frac{A}{\sqrt{1+\omega^2\tau^2}} \cos(\omega t - \theta) \tag{3.26}$$

Similarly,

$$V_{OUT-h}(t) = Le^{-\frac{t}{\tau}} \tag{3.27}$$

However the constant L, is still unknown. To find it, an auxillary condition is required: $V_{OUT} = V_{CO}$. Hence,

$$V_{CO} = L + \frac{A}{\sqrt{1+\omega^2\tau^2}}\cos(-\theta)$$

$$L = V_{CO} - \frac{A}{\sqrt{1+\omega^2\tau^2}} \tag{3.28}$$

and

$$V_{OUT}(t) = \left(V_{CO} - \frac{A}{\sqrt{1+\omega^2\tau^2}}\right)e^{-\frac{t}{\tau}} + \frac{A}{\sqrt{1+\omega^2\tau^2}}\cos(\omega t - \theta) \tag{3.29}$$

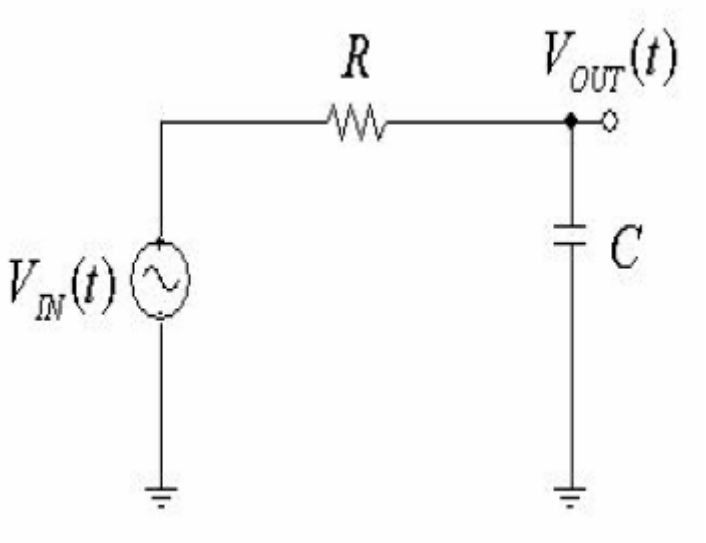


Figure 3.7: Sampling Network Model

As simple as this system is, with perfectly linear elements, it is still nonlinear. Closely examining the auxiliary condition explains the reason. First assume that the auxiliary

condition is zero; applying the definition of linearity (i.e. associativity and commutativity) it can be shown that indeed the system is linear for this special case. However, doing the same for any arbitrary initial condition yields a nonlinear system. This is a significant point because, if the sampling circuit is inherently nonlinear to begin with, linearizing it is more difficult. In fact all sampling circuits that do not reset the sampling capacitor are nonlinear. To linearize the circuit, τ must be minimized sufficiently to incrementally meet the desired linearity.

Even assuming a linear circuit (i.e. $V_{co} = 0$), in transient conditions, spectral purity is not implied. If the system has not yet reached a steady state, a pure sinusoid in will not result in a pure sinusoid at the output. Again, τ must be minimized to limit this effect, resulting in a circuit with a faster response that reaches steady state sooner. For a high-speed S/H this is a critical point because there is very little time to reach steady state conditions.

Assume that the sampling time duration equals $\kappa\tau$ with κ measuring the number of time constants required to attenuate the exponential term the amount SFDR (decibels) below the maximum acquired signal at this frequency. Since $\omega^2\tau^2 << 1$ and A = 1,

$$\kappa > \frac{SFDR}{20}\ln 10 + \ln 2$$

Each 6dB change in SFDR thus requires 0.7 more time constants. For example, SFDR=90 dB requires $\kappa > 11$ but SFDR=105 requires $\kappa > 12.8$.

3.1.7 Improved Component and Charge Injection Linearity with a Transmission Gate Switch

A transmission gate switch of Figure 3.8 consists of a NMOS in parallel with a PMOS. The PMOS device requires the inverse of the clock signal to work in unison with its

NMOS counterpart. The transmission gate increases the input range while linearizing the on resistance and reducing nonlinearities created by C_{SB} and C_{DB}. From the small signal model, R_{DS-N} is in parallel with R_{DS-P}.[3] [4] Similarly, C_{SB-N} and C_{SB-P} in addition to C_{DB-N} and C_{DB-P} are in parallel. From Equations 16 and 18 C_{SB} and C_{DB} nonlinearly decrease with increasing V_{SB} and V_{DB}. When V_{SB} increases for the NMOS, it decreases by the same amount for the PMOS since their sources are tied together and their bulks are at constant potentials.

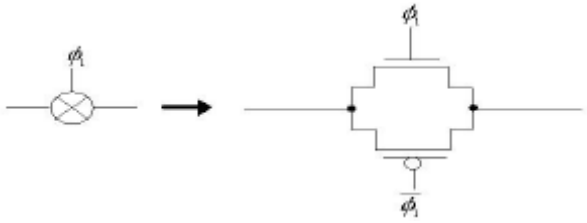


Figure 3.8: Transmission Gate

The parallel resistance of a TGATE in the linear region (assuming $V_{DS} = 0$) is

$$R_{\|} = (R_{DS-n} \| R_{DS-p}) = \frac{1}{\mu_n C_{OX} \frac{W_n}{L_n} (V_{GS-n} - V_{t-n} + \mu_p C_{OX} \frac{W_p}{L_p} (|V_{GS-p}| - |V_{t-p}|)}$$

But $V_{GS-n} = V_{DD} - V_{IN}$ and $V_{GS-p} = V_{IN}$. Thus

$$R_{\|} = (R_{DS-n} \| R_{DS-p}) \tag{3.30}$$

$$= \left[\mu_n C_{OX} \frac{W_n}{L_n} (V_{DD} - V_{t-n}) - \mu_p C_{OX} \frac{W_p}{L_p} |V_{t-p}| + V_{IN} (\mu_n C_{OX} \frac{W_n}{L_n} - \mu_p C_{OX} \frac{W_p}{L_p}) \right]$$

$R_{\|}$ is input independent ("linear") if the term with input dependent term in Equation is zero

$$\mu_n C_{OX} \frac{W_n}{L_n} - \mu_p C_{OX} \frac{W_p}{L_p} = 0$$

This can be accomplished by sizing the two devices appropriately. Furthermore, if $V_{t-n} = |V_{t-p}|$ and the oxide capacitances are equal, Equation 3.30 becomes

$$R_{\|} = \frac{1}{(\frac{W}{L}\mu)_{n,p} C_{OX} V_{DD}}$$

While perfectly linearized resistance is possible if the threshold voltages are constant and the devices have exactly equal conductance, the actual degree of linearization is considerably less. Assuming an input range from 1.5V to 3.5V and a $\Delta\gamma$ of $0.1\sqrt{V}$ between NMOS and PMOS, the change in the resistance due to threshold voltages alone is about 0.5%, equivalent to about 7.5 bits. Similarly, if the conductance mismatch is 2.5%, the resistance change due to this error alone is about 0.5%. Clearly, even with well matched NMOS and PMOS device characteristics, 9-bits of linearity appear to be the best case scenario. Recall however that this nonlinearity only becomes significant for the tracking of high-frequency inputs and the cancellation of charge injection as the clock frequency increases.

A similar analysis yields the nonlinear component of the total parasitic capacitance:

$$C_{\|} = C_{db-n} + C_{sb-n} + C_{db-p} + C_{sb-p}$$

$$C_{db-n} + C_{sb-n} = \frac{C_{jo-n}(A_{D-n} + A_{S-n})}{\sqrt{1 + \frac{V_{IN}}{\phi_{0-n}}}} + \frac{C_{jo-CH-n} A_{CH-n}}{\sqrt{1 + \frac{V_{IN}}{\phi_{0-n}}}} + \frac{C_{jo-SW-n}(P_{S-n} + P_{d-n})}{\sqrt{1 + \frac{V_{IN}}{\phi_{0-n}}}} \tag{3.31}$$

and

$$C_{db-p}+C_{sb-p}=\frac{C_{jo-n}(A_{D-p}+A_{S-p})}{\sqrt{1+\frac{V_{IN}}{\phi_{0-p}}}}+\frac{C_{jo-CH-p}A_{CH-p}}{\sqrt{1+\frac{V_{IN}}{\phi_{0-p}}}}+\frac{C_{jo-SW-p}(P_{S-p}+P_{d-p})}{\sqrt{1+\frac{V_{IN}}{\phi_{0-p}}}} \tag{3.32}$$

A partial cancellation of the nonlinear terms is possible for typical junction profiles if the NMOS and PMOS area and perimeters are adjusted appropriately. Because of the two square root terms, complete input dependence is not possible unlike the source-drain resistance. Consider a typical example to see the degree of linearization possible. The nonlinear cancellation is pairwise such that the first NMOS term acts to cancel the first PMOS term, continuing to the second and third terms in (3.31) and (3.32). Assuming equal numerators (C_{j0} Area product) and setting $\phi_{0-n}=\phi_{0-p}=0{:}65\text{V}$, VDD = 5V , $V_{IN-\max}=3{:}5\text{V}$, and $V_{IN-\min}=1{:}5\text{V}$, the total percentage change of either the NMOS or PMOS alone would be about 31% over the full input range, but reduces to 4% for the parallel combination. Hence, improvement by factor of 7.5 is achieved. In practice, computer simulation may be used to determine the device sizes that minimize the combined nonlinearity. Figure 3.9 illustrates these two component linearization effects. Estimating $\mu_n C_{OX}=75\frac{\mu A}{V}, \mu_p C_{OX}=25\frac{\mu A}{V}$ and $V_{tn}=|V_{tp}|=1V$ and letting $\left(\frac{W}{L}\right)_n=\frac{1}{3}\left(\frac{W}{L}\right)_p=\frac{50\mu m}{1\mu m}$ with $V_{CC}=5V$. R_{DSn} and R_{DSp} may be calculated for V_{IN} =0 to $V_{IN}=V_{CC}$. The parallel combination of R_{DSn} and R_{DSp} is linear between 1 and 4 volts. Similarly, taking CSB as an example of non-linear capacitances, assume C_{Sb-0n} and $C_{Sb-0p}=100\text{fF}$ for simplicity

The NMOS and PMOS bulks are connected to ground and V_{CC} respectively. Clearly, the parallel combination improves the linearity, especially around $\frac{1}{2}V_{CC}$ but has more overall capacitance than either individually.

Transmission gates also have a positive effect on charge injection. Because NMOS and PMOS devices have opposite charges in their channels, when they are turned off, positive and negative charge combine. The net effect is a reduction the magnitude of the charge injected onto the sampling switch. However, this is strictly effective only for one input input voltage level since the channel charges are opposite functions of VIN.

Transmission gates have four beneficial qualities that all improve the performance of either a NMOS or PMOS individually. These are increased operating range, reduced and linearized on resistance, reduced charge injection and linearized parasitic capacitance. Unfortunately, they cannot be optimally achieved under the same conditions. Consequently, a choice must be made as to which one is the most important for a particular design. The use of charge pumps to "fix" the VGS of each device removes the constraint on RDS, allowing an extra degree of freedom for the other three.

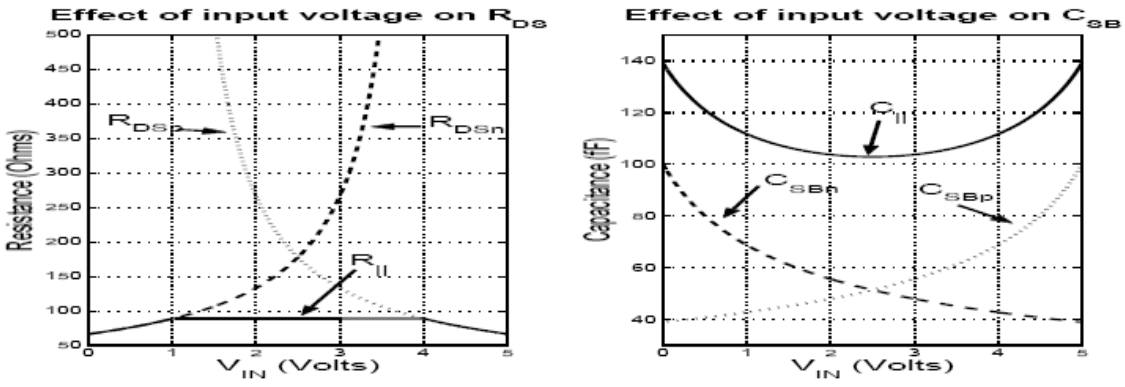


Figure 3.9: The linearizing effect of a transmission gate

3.1.8 Dummy Switch Compensation to Reduce Charge Injection

One method used to reduce charge injection is through the use of a dummy switch as shown in Figure 3.10 [6]. A transistor (Q2) is added with both its source and drain connected to the sampling node driven by a clock that turns on just as the sampling clock

turns off. If the devices are sized just right (generally $\left(\frac{W}{L}\right)_1 = \frac{1}{2}\left(\frac{W}{L}\right)_2$ is chosen to match Cgd_{-1} to $Cgd_{-2} + Cgs_{-2}$,Q2 will be able to absorb the charge injected onto CH. Charge distribution is a function of the clock edge as well as the impedances at the drain and source. Significant cancellation is not realistically attainable because of the difficulty of precisely modeling charge injection and the finite slopes of the clock edges. Realistically, this technique can reduce the amount of charge injected to about one fifth the level without Q2, which represents a significant improvement.

In fully differential circuits, as long as the injected charge is not input dependent and equal on both paths, no error will occur at the differential output. So, dummy Switch compensation is only a viable option for single ended S/Hs.

.

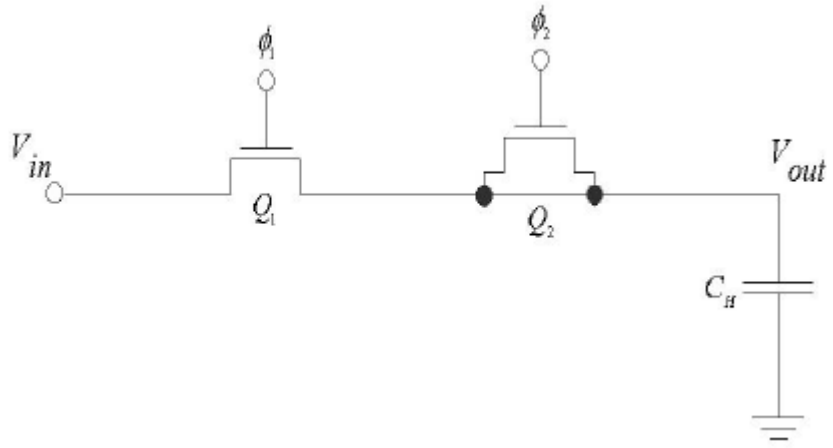


Figure 3.10: Compensating Charge Injection with a Dummy Switch

3.1.9 Summary of MOS Sampling Nonlinearities

Nonlinearities in MOS sampling circuits are primarily due to components (e.g. R_{DS} , C_{Sb} and C_{db}), clock effects, charge injection and the sampling network. The contradictory requirements to minimize them are listed below.

Drain-Source Resistance pumping must be done in a manner that keeps $V_{GS} - V_t$ constant to keep RDS input independent. Equal conductance NMOS and PMOS devices in a TGATE also do so but are limited by threshold voltage and conductance matching.

Parasitic Capacitance (i.e. C_{Sb} and C_{db}) must be minimized, leading to smaller device sizes. A TGATE with approximately equally sized NMOS and PMOS devices tends to counteract the nonlinearity compared to either individually.

Clock Effects require some form of series sampling to be used to remove input dependence on the sampling time. Minimizing the full-scale swing and maximizing the supply voltage are beneficial. Faster turn off times also reduce this effect.

Charge Injection needs a minimum length, bottom plate sampling, maximum pump overdrive, minimum bias resistance and larger full-scale swing to reduce input dependence.

Sampling Network nonlinearity requires a smaller time constant. Therefore L must be minimized and W maximized to reduce the sampling switch resistance. Pumping with maximum overdrive will also decrease the input switch on resistance.

Clearly, there are competing requirements and compromises to be made. The most obvious trade off involves the maximum supply voltage, overdrive and headroom versus minimum channel length (as T_{OX} is scaled with $L_{\min}$). It is best to proceed by considering each error source individually; through simulation and experimentation compare the relative magnitudes of each effect after which the designer find the best compromise for the given design goals.

Chapter 4

CMOS Sample and hold Circuits

Each sampling circuit discussed in the previous section is incapable of driving a load. Consequently, a buffer of some kind is needed at the output. An open loop approach is inherently fast, but cannot achieve the same linearity as closed loop negative feedback architectures. A typical open loop buffer is simply a follower of some kind. Closed loop implementations must maintain stability while pushing the limits on open loop gain and unity gain frequency in order to be practical for high-speed high-resolution S/Hs. The large gain requirements in closed loop amplifiers dictate the use of operational amplifiers. Therefore, the options are either to maximize the speed of a closed loop design or the linearity of open loop architecture.

Some S/H circuits require input buffers while others do not. Justification for this lies at the system level. For example, a S/H that follows a filter of some kind must not alter the transfer function of the filter. Although it is possible to design the filter and S/H input network together, design complexity is greatly increased since neither is a trivial undertaking and now extra constraints and therefore compromises are introduced.

4.1 Single Ended Versus Differential Operation

Among all the analog circuits, Operational Amplifier (op amp) is perhaps the most fundamental and commonly used one, which is usually an integral part of other analog and mixed signal systems. CMOS op-amps often have a lower transconductance for a given current, lower gain, slower speed, higher input-referred offset and input-referred noise voltage than their BJT counterparts[24]. However, analog and digital circuit are now often integrated together -- signals that originate in analog form often need to be interfaced to the digital circuits in many complex systems, such as A/D and D/A converters. With CMOS technology becoming dominant because of its smaller size and

less power dissipation, the integration requirement of mixed signal systems provide a strong economic incentive to use CMOS op-amps.

Unlike conventional signal-ended op amps, fully differential op-amps have a differential input and produce a differential output. They have some advantages over the signal-ended op amps and are widely used in modern CMOS integrated circuits. For example, by generating "complementary" outputs on both differential arms, fully differential op amps provide roughly twice output swing as their single-ended counterparts. And achieving a large swing becomes more and more important in the modern CMOS circuit design as the supply voltage decreases. Besides, as a balanced circuit with perfectly symmetric components on each side, fully differential op-amps are less susceptible to common-mode noise and even-order nonlinearities. A disadvantage of this topology is that it usually needs a common-mode feedback circuit to control the common-mode output voltage, which increases the design complexity.

Fundamentally, fully differential circuits perform better than their single ended counterparts. As stated above the advantage lies in the fact that common mode effects and even order nonlinearities are cancelled in symmetric fully differential circuits [24]. Consider common mode effects, when the differential output, the difference of the two outputs, is taken they cancel each other out. With individual differential inputs and outputs V_{ip}, V_{im}, V_{op} and V_{om} the combined differential input and output are

$$V_{id} = V_{ip} - V_{im} \tag{4.1}$$

and

$$V_{od} = V_{op} - V_{om} \tag{4.2}$$

Hence, if a differential signal, $A\sin(\omega_1 t)$, was applied along with a common mode signal, $B\cos(\omega_2 t)$, and the outputs simply equal the inputs then,

$$
\begin{aligned}
V_{od} &= V_{id} \\
&= V_{ip} - V_{im} \\
&= \frac{1}{2}\left(A\sin(\omega_1 t) + B\cos(\omega_2 t) - \frac{1}{2}\left(-A\sin(\omega_1 t)\right) + B\cos(\omega_2 = t)\right) \\
&= A\sin(\omega_1 t)
\end{aligned}
\tag{4.3}
$$

What (4.3) shows is that if there is perfect matching between the the two channels, no error common to both channels is present at the output.

If there is some mismatch between the two channels, then the difference of the mismatch shows up at the output. For example, if $V_{op} = V_{ip}$ and $V_{om} = (1+\delta)V_{im}$ where delta is the mismatch factor, then,

$$
\begin{aligned}
V_{od} &= V_{op} - V_{om} \\
&= V_{ip} - (1+\delta)V_{im} \\
&= \frac{1}{2}\left(A\sin(\omega_1 t) + B\cos(\omega_2 t)\right) - \frac{1+\delta}{2}\left(-A\sin(\omega_1 t) + B\cos(\omega_2 t)\right) \\
&= A\sin(\omega_1 t) + \frac{\delta}{2}\left(A\sin(\omega_1 t) - B\cos(\omega_2 t)\right)
\end{aligned}
\tag{4.4}
$$

So, the differential output is the sum of the ideal output and an error term scaled by the mismatch factor δ. If the mismatch is constant, then it manifests itself as either a gain error or an offset. However, if the mismatch is a function of input, then distortion is introduced.

Next consider the effect of nonlinearities [14]. Assume that the system is described by the following equation

$$
y = \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 \tag{4.5}
$$

(4.5) expresses the system in terms of a linear term, an even order nonlinear term and an odd order nonlinear term. Given differential inputs

$$x_1 = A\cos(\omega t) \tag{4.6}$$

and

$$x_2 = -A\cos(\omega t) \tag{4.7}$$

the differential outputs are then

$$y_1 = \frac{\alpha_2 A^2}{2} + \left(\alpha_1 A + \frac{3\alpha_3 A^3}{4}\right)\cos(\omega t) + \frac{\alpha_2 A^2}{2}\cos(2\omega t) + \frac{\alpha_3 A^3}{4}\cos(3\omega t) \tag{4.8}$$

and

$$y_2 = \frac{\alpha_2 A^2}{2} - \left(\alpha_1 A + \frac{3\alpha_3 A^3}{4}\right)\cos(\omega t) + \frac{\alpha_2 A^2}{2}\cos(2\omega t) - \frac{\alpha_3 A^3}{4}\cos(3\omega t) \tag{4.9}$$

Consequently, the differential output is

$$y_1 - y_2 = \left(2\alpha_1 A + \frac{3\alpha_3 A^3}{4}\right)\cos(\omega t) + \frac{\alpha_3 A^3}{2}\cos(3\omega t) \tag{4.10}$$

Examining the differential output, no even order nonlinearity is present (noted by the absence of any α_2 or $\cos(2\omega t)$ terms). This analysis can be extended to higher order terms with a similar result. In a practical sense, differential operation relaxes the effect errors have on the system. Now, droop, charge injection or any other non-ideal characteristic of a differential S/H need only be input independent while not pushing the circuit out of its region of normal operation to not introduce distortion. For this reason, fully differential operation is requisite for a high-performance S/H.

4.2 Sample and Hold Configurations

The trade-offs studied above can be alleviated through innovations at both architecture and circuit level. In this section, we describe a number of conventional SHAs so as to illustrate their design challenges [2].

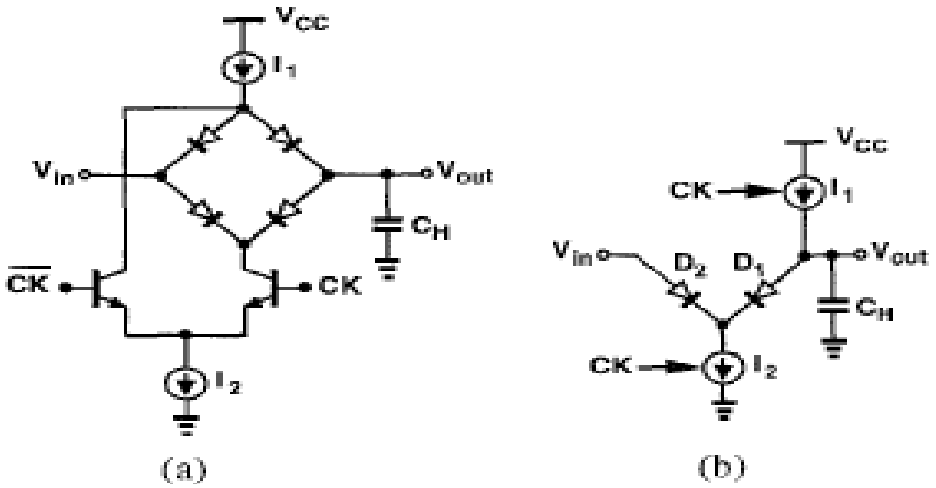


Figure 4.1: (a) Simplified diode bridge, (b) bridge with two top diodes removed and bottom current source converted to single-ended form,

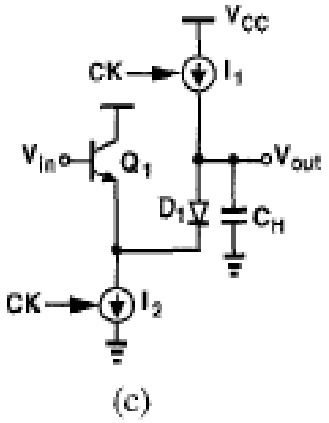


Figure 4.1: (c) Bridge with input follower to lower kickback noise.

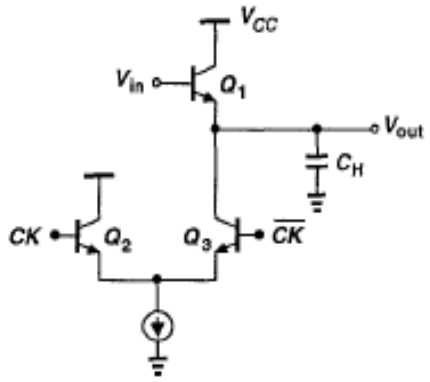


Figure 4.2: Switched emitter follower

In order to suppress input-dependent pedestal errors in a SHA, the sampling switch can be embedded in a feedback loop such that it always turns off with a constant gate-source voltage. Shown in Fig. 4.3, the conventional closed-loop

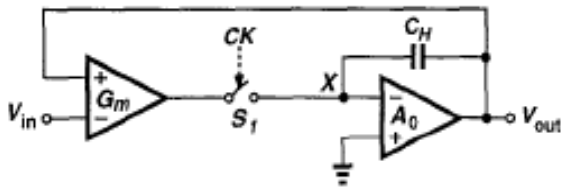


Figure 4.3: Closed-loop SHA architecture.

architecture incorporates this concept in a two-stage op amp topology. Since node *X* is a virtual ground, the charge injected by S_1 is relatively constant. However, stability and settling considerations that plague typical two-stage op amps limit the speed of this architecture.

A SHA configuration especially popular in CMOS technology is shown in Fig. 4.4(a). Based on the series sampling technique, this approach utilizes a virtual ground

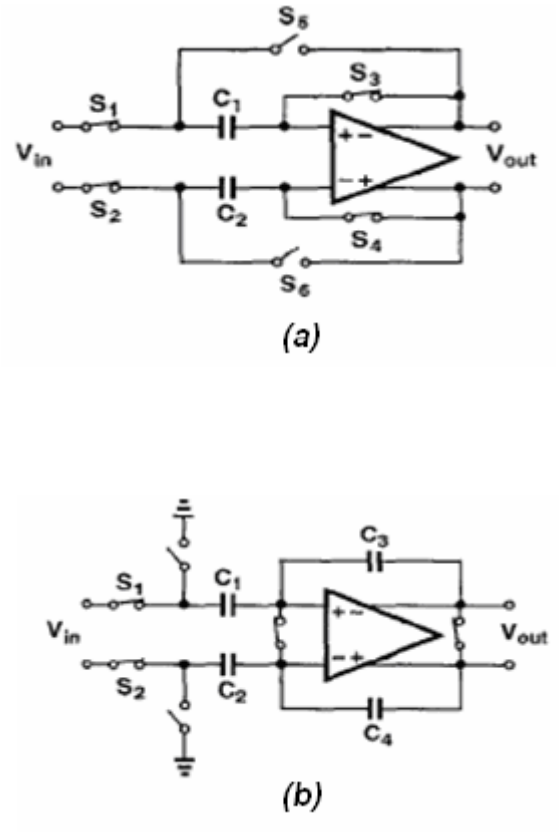


Fig 4.4: a) Unity-gain amplifier b) Alternate topology with independent Input and output CM levels

to perform a precise hold operation. Since at the end of the acquisition mode, S_3 and S_4 turn off before S_1 and S_2, the charge injected by the former two is constant and equal, disturbing only the output CM level. The challenge, however, lies in the design of the op amp. High-speed applications mandate the

use of a telescopic cascode op amp for fast settling, but it is difficult to short the input and the output of such a topology. For this reason, the configuration in Fig. 4.4(b) proves useful because it allows independent choice of the input and output CM levels. Nevertheless, the trade-offs among speed, gain, and output voltage swing of op amps and the issues mentioned above in relation with MOS switches make this approach less attractive at low supply voltages.

Fig. 4.5 depicts a SHA architecture originally proposed by Ryan [8] and later modified by Petschacher et al. [17]. Employing transconductance stages G_{m1} and G_{m2} and transresistance

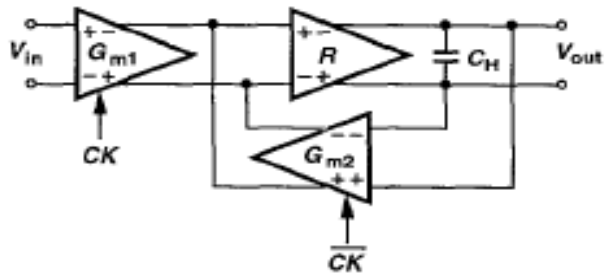


Figure 4.5: SHA architecture with multiplexed input

stage R (with $G_{m1}R = G_{m2}R = 1$), the circuit operates as follows. In the acquisition mode, Gm1 and R function as a unity-gain amplifier, allowing Vout to track Vin. In the hold mode G_{m1} is disabled, G_{m2} is enabled, and G_{m2} and R are configured as a unity-gain amplifier, thereby retaining the sampled value of V, across CH.

The primary drawback of this circuit is the linearity required of the open-loop amplifier $G_{m1}R$. Furthermore, the deviation of $G_{m2}R$ from unity and the drive capability of the R stage are also problematic. With 5-V supplies, various correction techniques can be applied to a basic linearized differential pair to achieve IO-bit linearity at sampling rates as high 75 MHz [8], but with a 3-V supply the design becomes much more difficult.

Another SHA topology in which the sampling switch introduces only a constant pedestal is the recycling circuit shown in Fig. 4.6 [18]. In the sampling mode, S_1 & S_2 and S_4 are on and

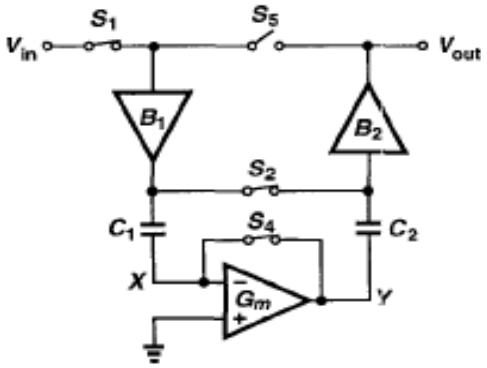


Fig 4.6: Recycling SHA architecture.

the G, stage establishes a virtual ground at X and *Y*. In the transition to the hold mode, first S_4 turns off, subsequently S_1 and S_2 turnoff and S_5 turn on. Thus B1, Gm and B2 constitute a unity gain feedback loop, holding the sampled level on C1. The settling speed of this circuit is still limited owing to the poles contributed by the three stages in the hold-mode feedback loop.

Another popular configuration is double sampling configuration as shown in fig 4.7. Like the name implies, double-sampling doubles the sample rate. This is possible because the flip around and switched capacitor amplifier S/Hs each have a reset phase that takes approximately half the sampling period. By eliminating the reset phase of the amplifier and adding a second set of sampling capacitors and switches, the throughput is doubled. A single ended implementation (for simplicity) of a flip around S/H is shown in

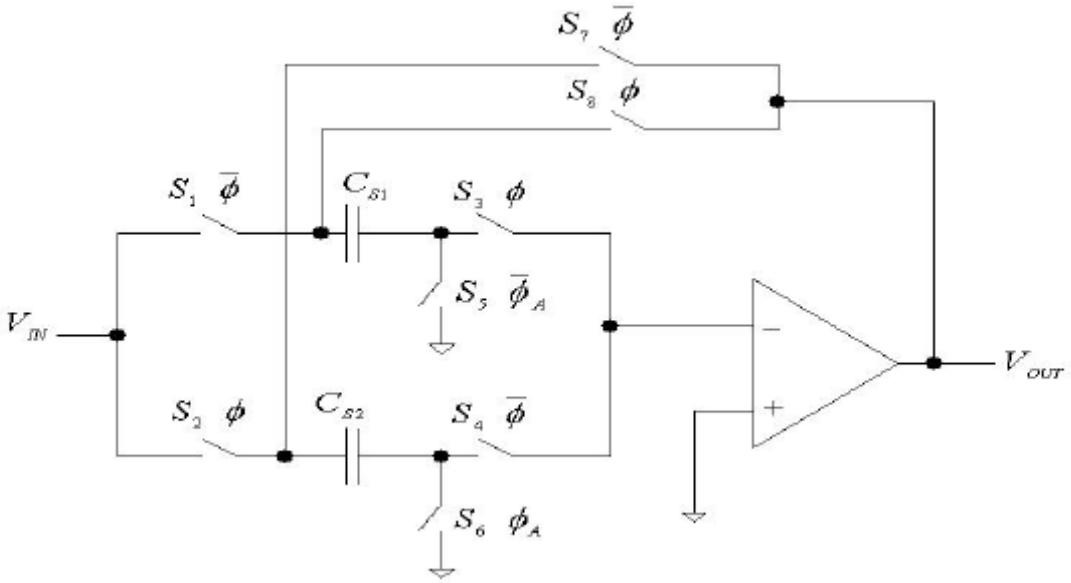


Figure 4.7: Double-Sampled S/H

Figure 4.7 [16]. On ϕ, S_2, S_3, S6 and S8 are on with S_1, S_4, S_5 and S_7 off. This places C_{S2} in open loop in sample mode and C_{S1} in the Op-Amp's feedback path, forcing its output to the held voltage on it. Next, ϕ_A turns off, sampling the input on C_{S2}. Shortly thereafter the clocks change phase. Now ϕ is high. Consequently, S_1, S_4, S_5 and S_7 are on and S_2, S_3, S_6 and S_8 are off, the reverse of the previous situation.

Error sources for this circuit are the same as outlined for the flip around architecture with two significant additions. Since there are now two paths, they must match or else errors will result, resulting in amplitude modulation at the output. Caused by non-uniform sampling, timing skew is the most significant new errors. Timing skew may be removed for example by adding an additional switch between S_5, S_6 and ground. This new switch will operate at the full sampling rate of the converter, twice as fast as the other clocks, and be responsible for taking the samples. Using two OTAs is not a viable option because the offsets and gains must match. Even with only one OTA, nonlinearity is introduced by a memory effect at the OTA input; the input capacitance of the OTA stores the previous sample on it, detrimentally affecting the subsequent sample.

The effect of these nonlinearities is large enough to reduce SFDR, typically, to well below 80dB.

In order to achieve large signal linearity improvements, it will be necessary to invent new double-sampling circuit techniques.

Chapter 5

Design of Sample and Hold Circuit

5.1 Introduction

In this chapter the architecture to be designed is explained and the various design steps for the designing of each of the sub modules has been explained.

After design the results of the S/H circuits are shown which will include the curve of the gain and phase margin of the op-amp and the sample response of the circuit.

5.2 Architecture

In order to maximize the speed of a closed loop S/H, the fully differential unity gain sampler has been realized as shown in the figure 5.1[2] [4]

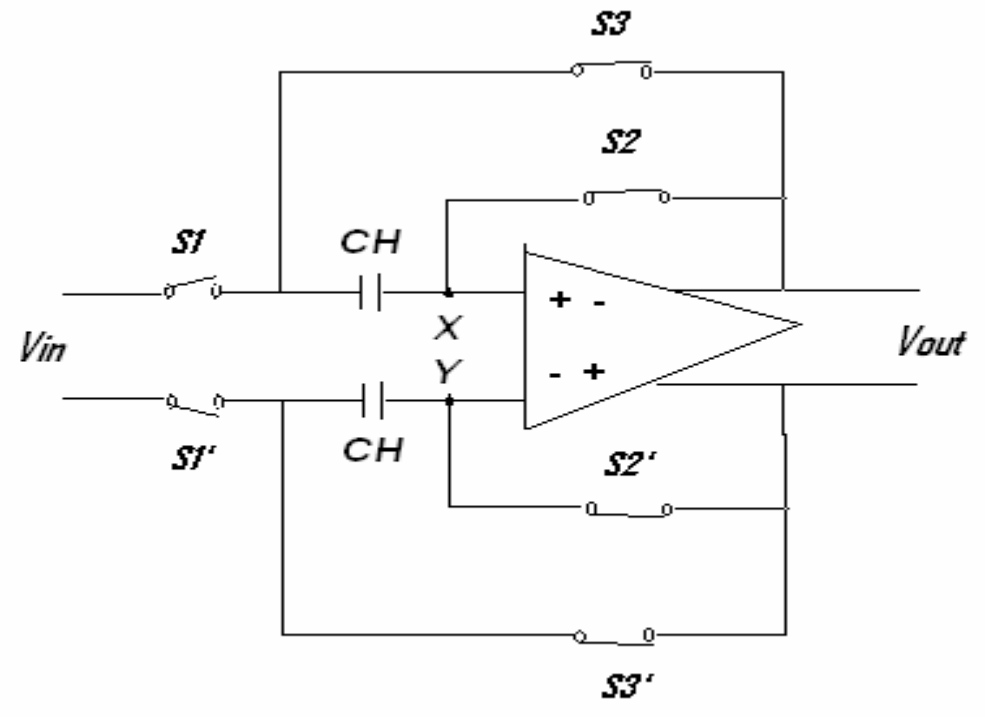
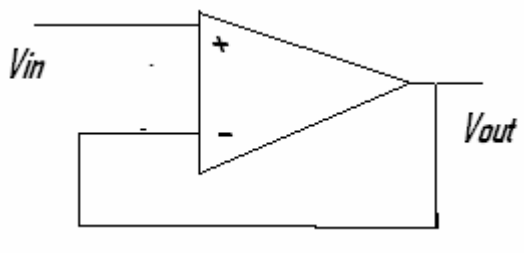


Figure 5.1: Differential unity gain sampler

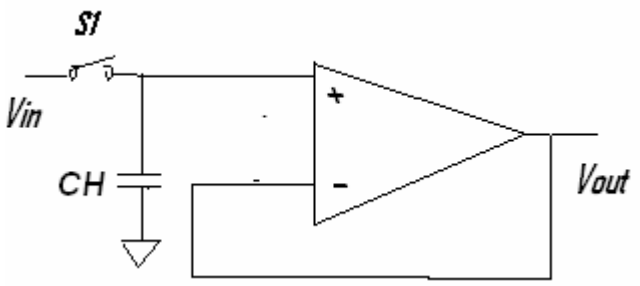
It has the benefits of open loop sampling and a closed loop hold with an OTA in unity gain negative feedback to buffer and hold the sampled voltage while maximizing the limited gain bandwidth product of the OTA. It operates on two-phase non-overlapping clocks with a duty cycle of approximately 50% to divide settling time evenly between the sampling network and OTA. Departure from 50% duty cycle to favor the OTA is not worth the added complexity and performance degradation of the clock generation circuitry.

5.2.1 Unity Gain Sampler

While a unity gain sampler can be realized with no resistor or capacitor in the feedback network (fig 5.1(a)), for discrete time applications, it still requires a sampling circuit. We may therefore conceive that the circuit shown in fig 5.2 as a sampler/buffer [4] [24]. However, the input-dependent charge injection by S1 and CH limits the accuracy here.



a)

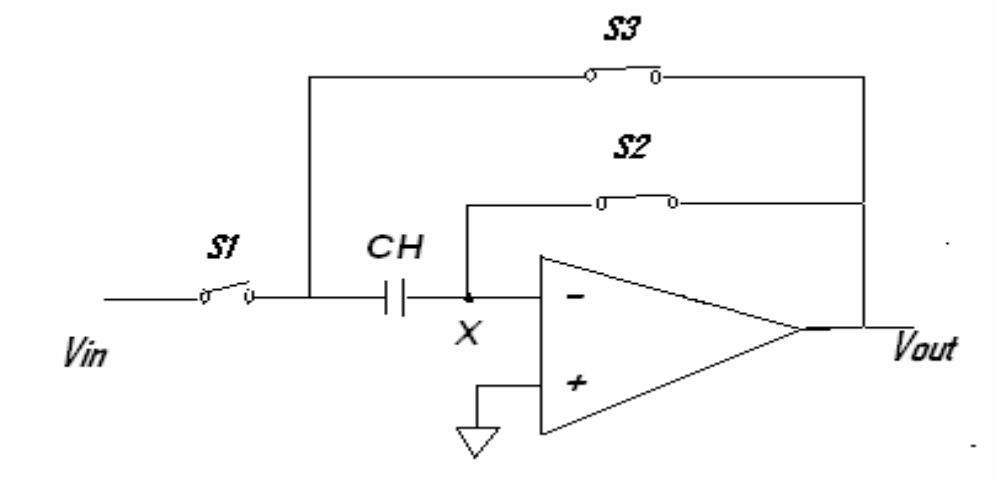


b)

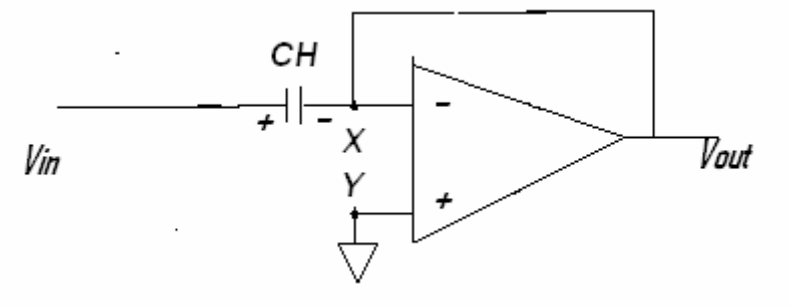
Figure 5.2: a) Unity-gain buffer, b) sampling circuit followed by unity-gain buffer

Now consider the topology depicted in figure 5.3, where three switches control the sampling and amplification mode. Here, with the proper control timings of clocks at the

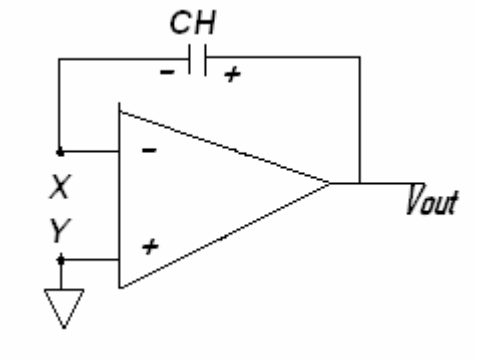
switches, the charge injected by S1 and S3 is unimportant and the channel charge of S2 results in a constant offset voltage.



a)



b)



C)

Figure 5.3: a) Unity gain sampler, b) circuit in sampling mode, c) circuit in amplification mode

The input dependent nature of the charge injected by the switches allows complete cancellation by the differential operation shown in figure 5.1.

5.3 Design Specification

Table 5.1 below shows the complete specification of the sample and hold circuit which is being achieved in this work.

Table 5.1: Specifications of Sample and Hold circuit of figure 5.1

Supply Voltage	5Volts
Sampling Frequency	20Mhz
Input Frequency	5Mhz
Accuracy of Sample and Hold circuit	10-Bit
Power	Minimum(<2mW)
Input Range	0.5-3Volts
Gain Of Op-amp	70db
Bandwidth Requirement	200Mhz
Phase Margin	>45 degrees
Slew Rate	

5.4 Design Steps

The most important part of this sample and hold configuration is the designing of the operational amplifier.

Here the fully differential folded cascade operational amplifier has been employed for the successful development of the configuration.

Next sections describe the steps used for the designing of the operational amplifier and the Common mode feedback circuit.

5.4.1 Design of operational amplifier

In this section design procedure of the folded cascode configuration shown in Figure 5.4 will be stated.

1) This design procedure assumes that the gain at DC (A_v), Unity gain bandwidth (UGB). Input Common Mode Range (V_{in} min and V_{in} max), load capacitance (C_L), slew rate (SR), output voltage swing (V_{out} max and V_{out} min) given.

2) Choose the smallest device length that will keep the channel modulation parameter constant and give good matching for current mirrors.

3) From the slew rate and load capacitance specification the current in the two braches can be obtained using

$$I_7 = SR \times C_{L1,\,L2}$$

4) M3 and M4 should provide this current and the current through input transistors M1 and M2, accordingly their W/L ratios can be obtained.
   One can assume that the current provide by M3 and M4 will equally divide between the input transistors and output branches.

5) Using output swing the calculations of the bias voltage of M6, M7, M8, M9 can be done.

6) As we now know the current supplied by transistors M3 and M4, the current mirror can be designed for that current.

7) Twice the current of the input transistors will be supplied by the tail current transistor M5.

8) Design of tail current source and input transistors also depends upon the given minimum input voltage swing.

9) Gain specifications can be met by varying the length of the load transistors and Gm of the input transistors.

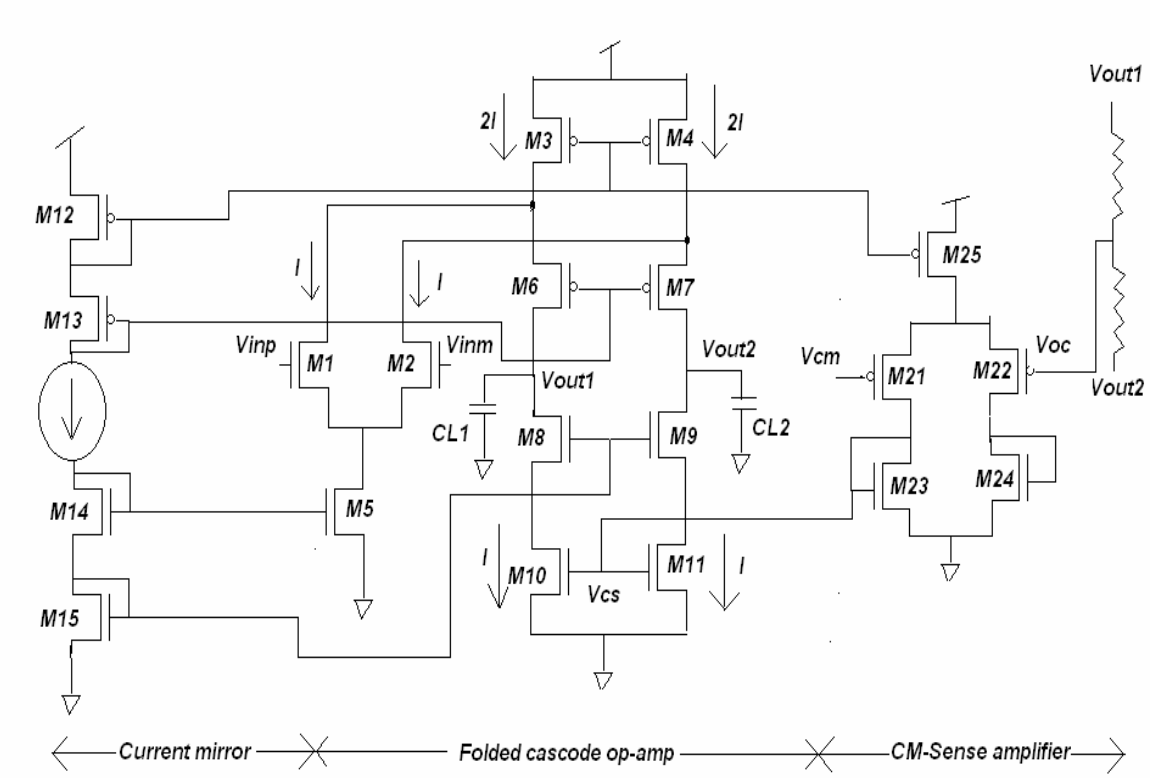


Figure 5.4 Fully Differential Folded Cascode Configuration

10) Bandwidth specifications can be met by varying the W/L ratios of the tail current or the input transistors. Our main aim here is to increase the current through input transistors.

11) In folded cascode design the load capacitances themselves act as compensating capacitances. So compensating folded cascode is quiet easy than the 2-stage op-amp [5] [21].

12) After folded cascode op-amp has been designed, now comes the designing of Common Mode Feedback Circuit which is necessary to maintained the output common mode voltage at a constant value,

The design procedure of CM-Sense amplifier is discussed in next section.

5.4.2 Design of Common mode feedback circuit

Common Mode Feedback (CMFB) is used to increase Common Mode Rejection Ration (CMRR), improve Input Common Mode Range (ICMR) and define output common mode level.

A straight forward way to detect the CM output level is to use two equal resistors, as shown in the figure 5.5 [24]. The voltage between the two resistors is

$$V_{oc} = \frac{V_{o1} + V_{o2}}{2} \tag{5.1}$$

This voltage is subtracted from the desired CM output voltage V_{CM}, and scaled by the differencing CM-sense amplifier in Figure 5.5(b) that consist of source-coupled pair $M_{21} - M_{22}$, diode connected loads M_{23} and M_{24} and the tail current source M_{25}.

The output of the amplifier, which drives the CMC input of the op-amp, is

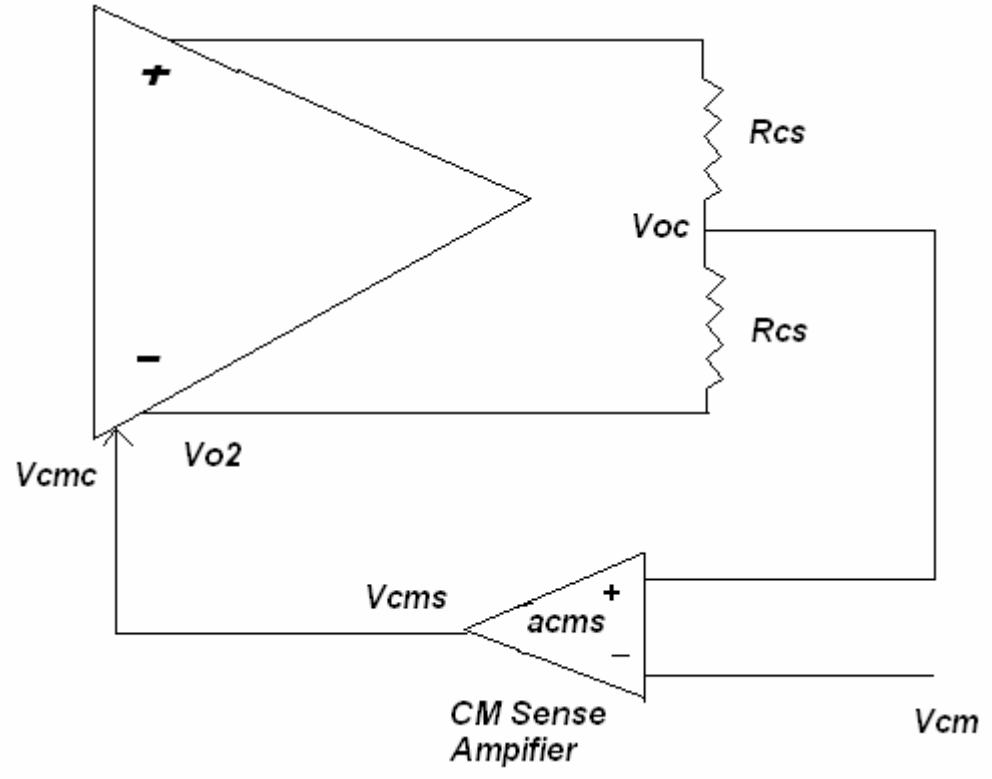
$$V_{cms} = a_{cms}(V_{oc} - V_{CM}) + V_{CSBIAS} \tag{5.2}$$

If $V_{oc} = V_{CM}$, $V_{cms} = V_{CSBIAS}$ .Therefore, for the circuit of Figure 5.4(b) $V_{CSBIAS} = V_{GS23} - V_{SS}$ when $I_{D23} = \frac{I_{D25}}{2}$. The value of V_{GS23} [or equivalently, I_{D23} and $(W/L)_{23}$] is chosen so that $I_{D23} = I_{D10} / I_{D11}$ when $V_{oc} = V_{CM}$.

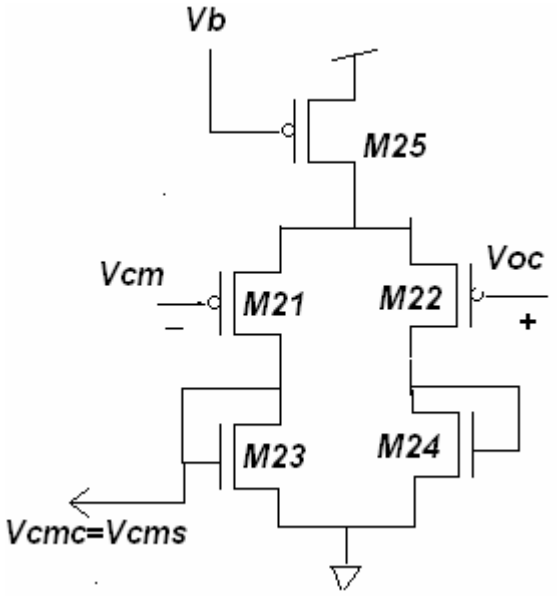
Here a_{cms} is the small signal voltage gain of the CM-sense amplifier which can be given as

$$a_{cms} = \frac{v_{cms}}{v_{oc}}\bigg|_{CMFBLoopGain} = \frac{1}{2}\frac{g_{m21}}{g_{m23}} \tag{5.3}$$

Here, we have assumed that the CM gain of the CM sense amplifier is much smaller in the magnitude that in Differential Mode gain. The factor of 1/2 multiplies $\frac{g_{m21}}{g_{m23}}$ because the output is taken from one side of the differential amplifier.



a)



b)

Figure 5.5: (a) CMFB using a resistor divider, (b) Schematic of CM-sense amplifier

5.4.3 Design of Sampling capacitor

The design of the sampling capacitor is based on *kT/C noise* (sometimes also called the *sampling noise)* .The noise on the resistance actually is the deciding factor of the value of the capacitor.

For example, in this design our accuracy is of 10-bit, that means we have a margin of $\pm$ 4mV around the exact value.

As discussed in section 3.1.5 we can use the below formula for calculating the value of the sampling capacitor

$$V^2 = \frac{KT}{C_{samp}}$$

or

$$V = \sqrt{\frac{KT}{C_{samp}}}$$

where K is the Boltzmann's constant ($1.38 \times 10{-}23$ JK−1), and T is the sampling clock period.

5.5 Results and Discussions

Using the above procedure in designing of sample and hold circuit, we have met the desired specifications quiet reasonably and will be shown later in the section.

Figure 5.6 below shows the schematic of the Sample and Hold circuit which make use of transmission gate as a switch and op-amp as the basic fundamental building block.

Figure 5.7 shows the output of the Sample and Hold circuit with input and sampled waveform.

Figure 5.8 shows the enlarged view of the sampled output.

Figure 5.9 shows the schematic of the fully differential folded cascade op-amp.

Finally, Figure 5.10 and 5.11 shows Gain and Phase Response of the Fully Differential folded Cascode Op-amp respectively.

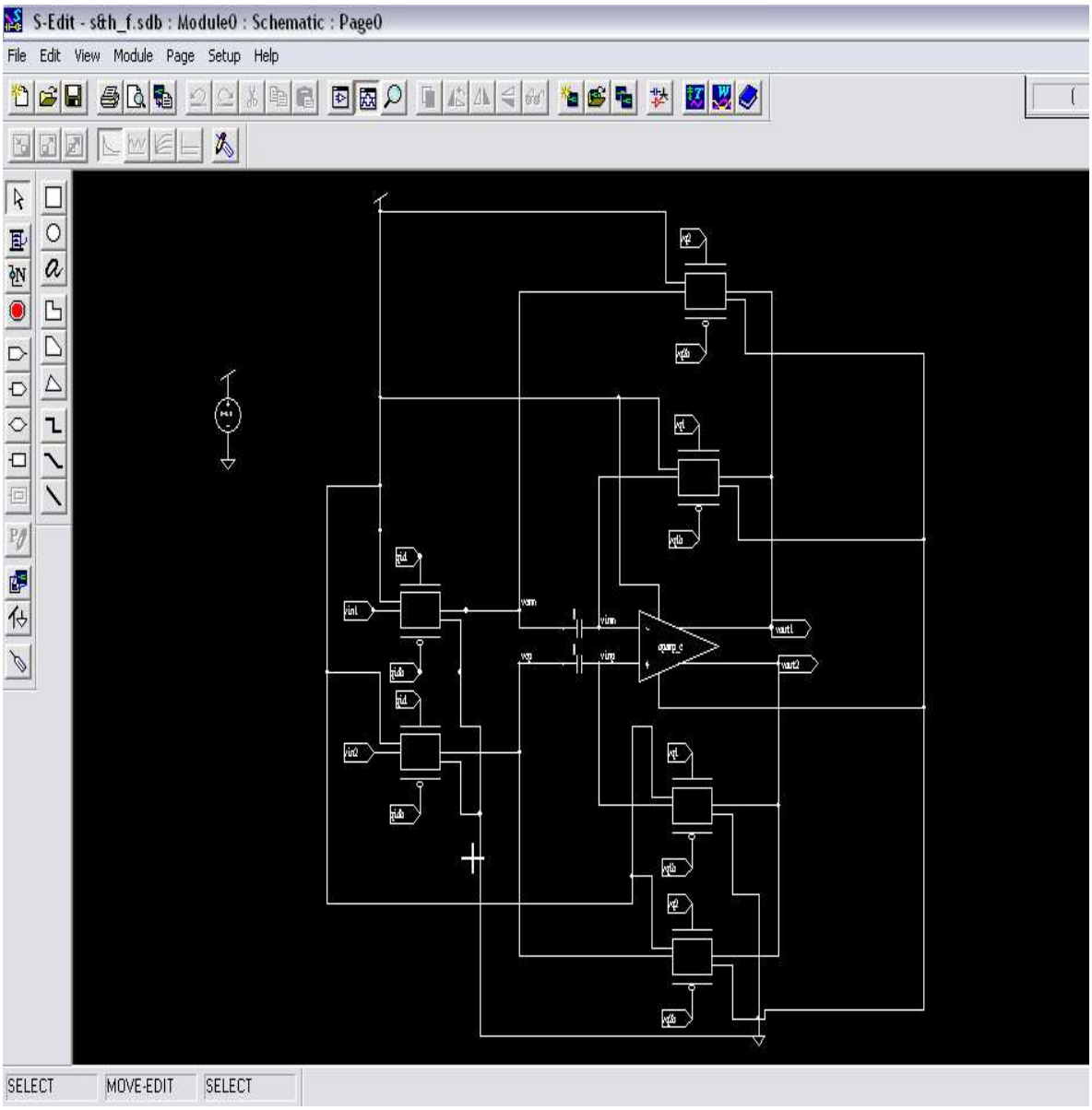


Figure 5.6: Schematic of the Sample and Hold circuit

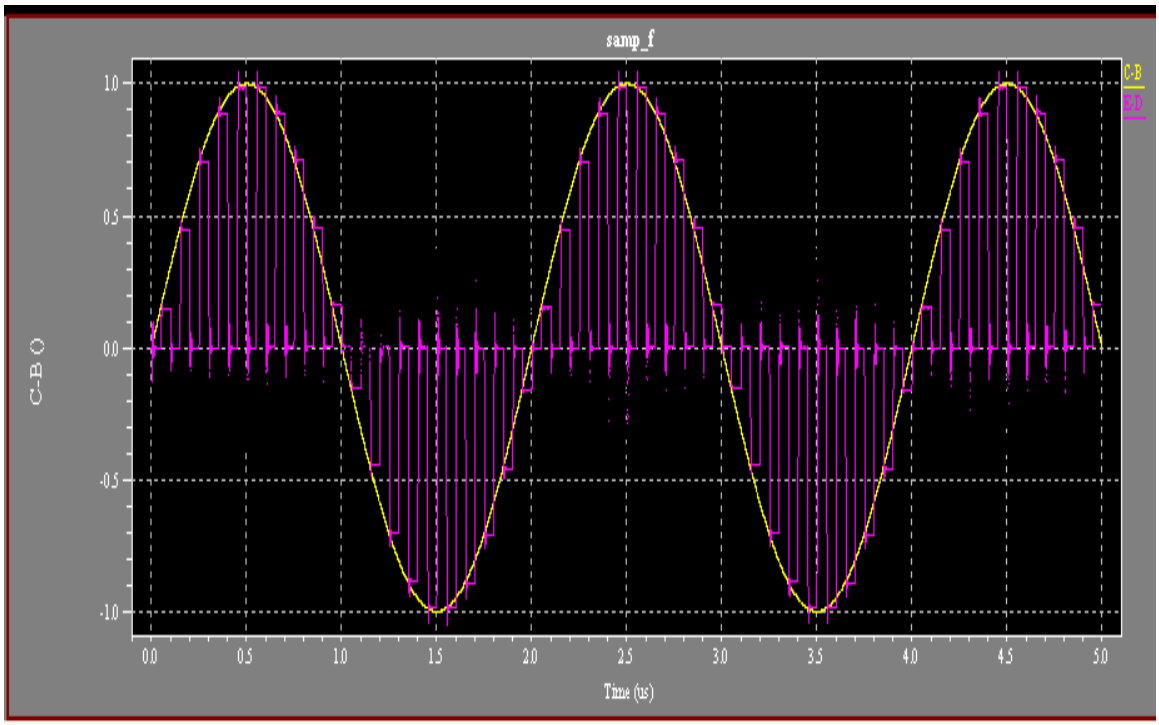


Figure 5.7: Output of the Sample and Hold circuit showing input and sampled waveform

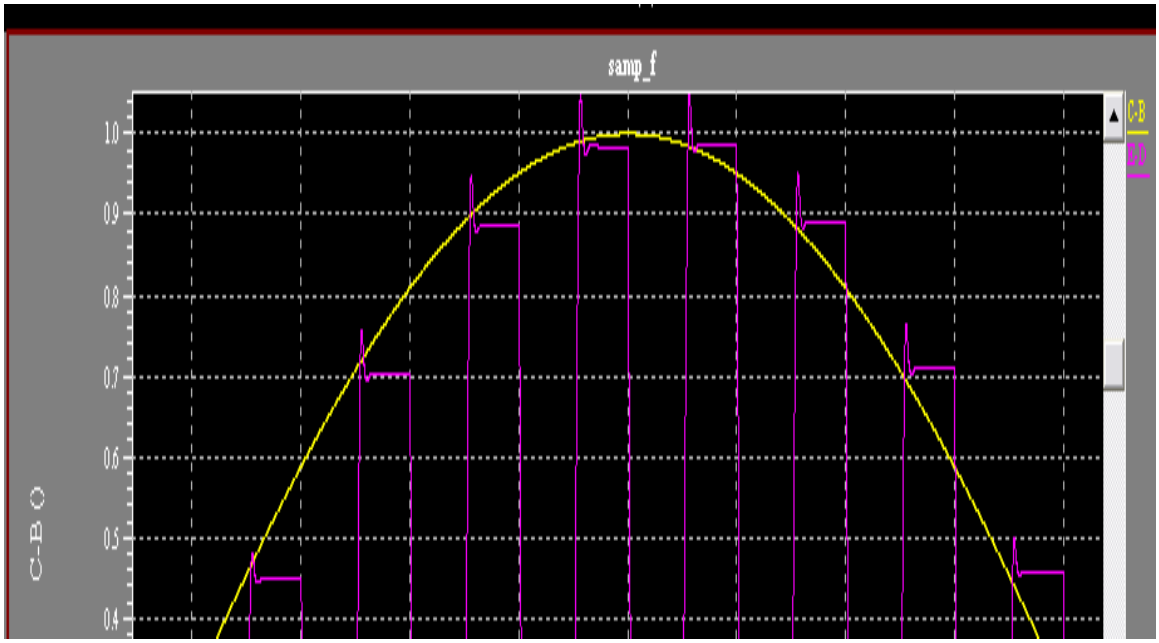


Figure 5.8: Showing enlarged sampled output

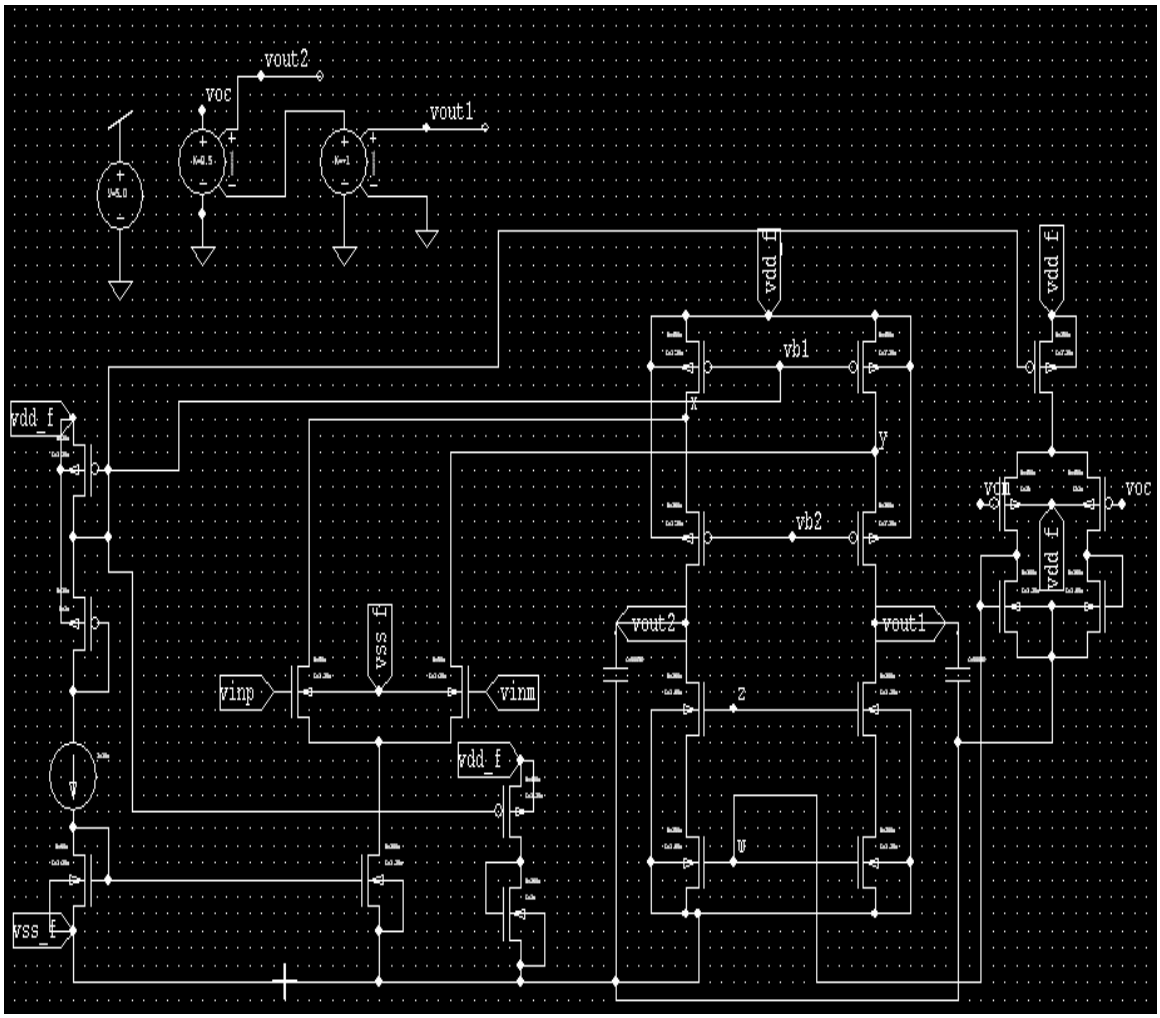


Figure 5.9 Schematic of Fully Differential Folded Cascode Op-amp in Tanner Tool

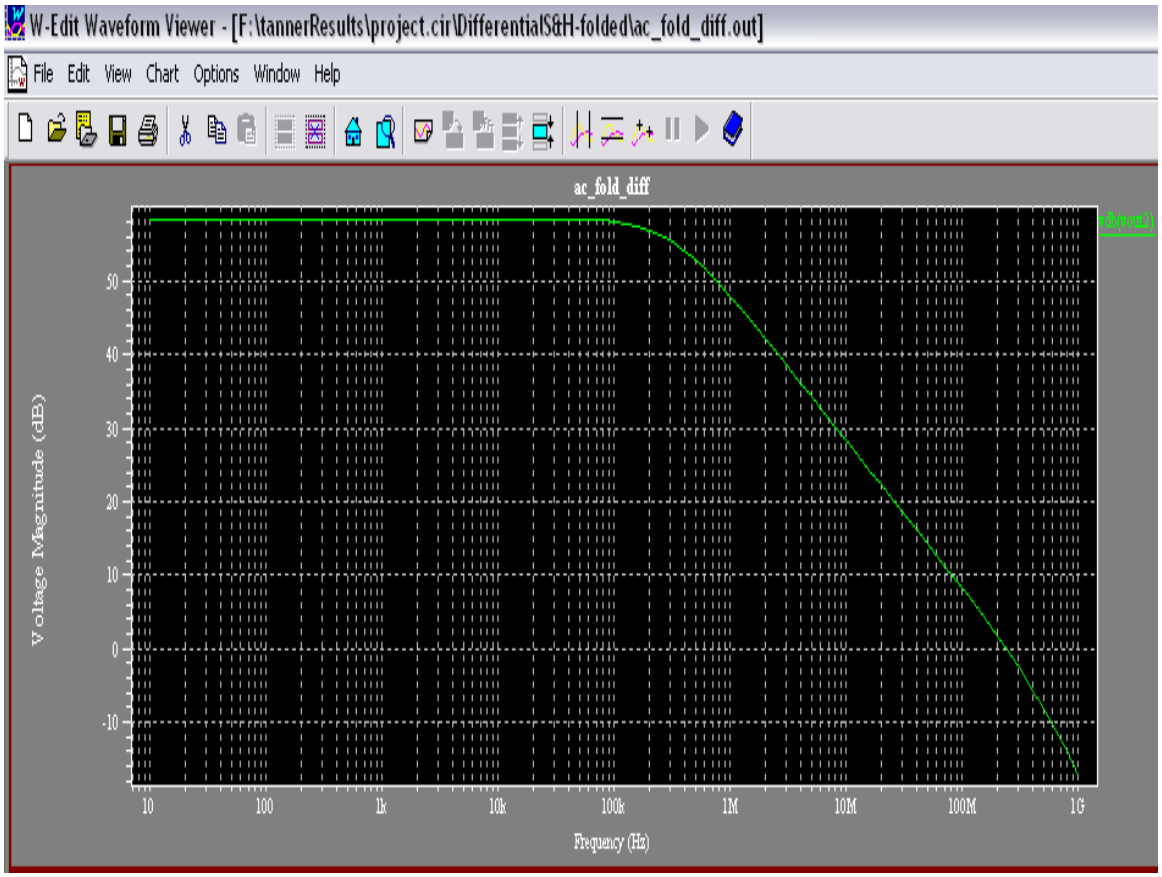


Figure 5.10: Gain Response of the Fully Differential folded Cascode Op-amp

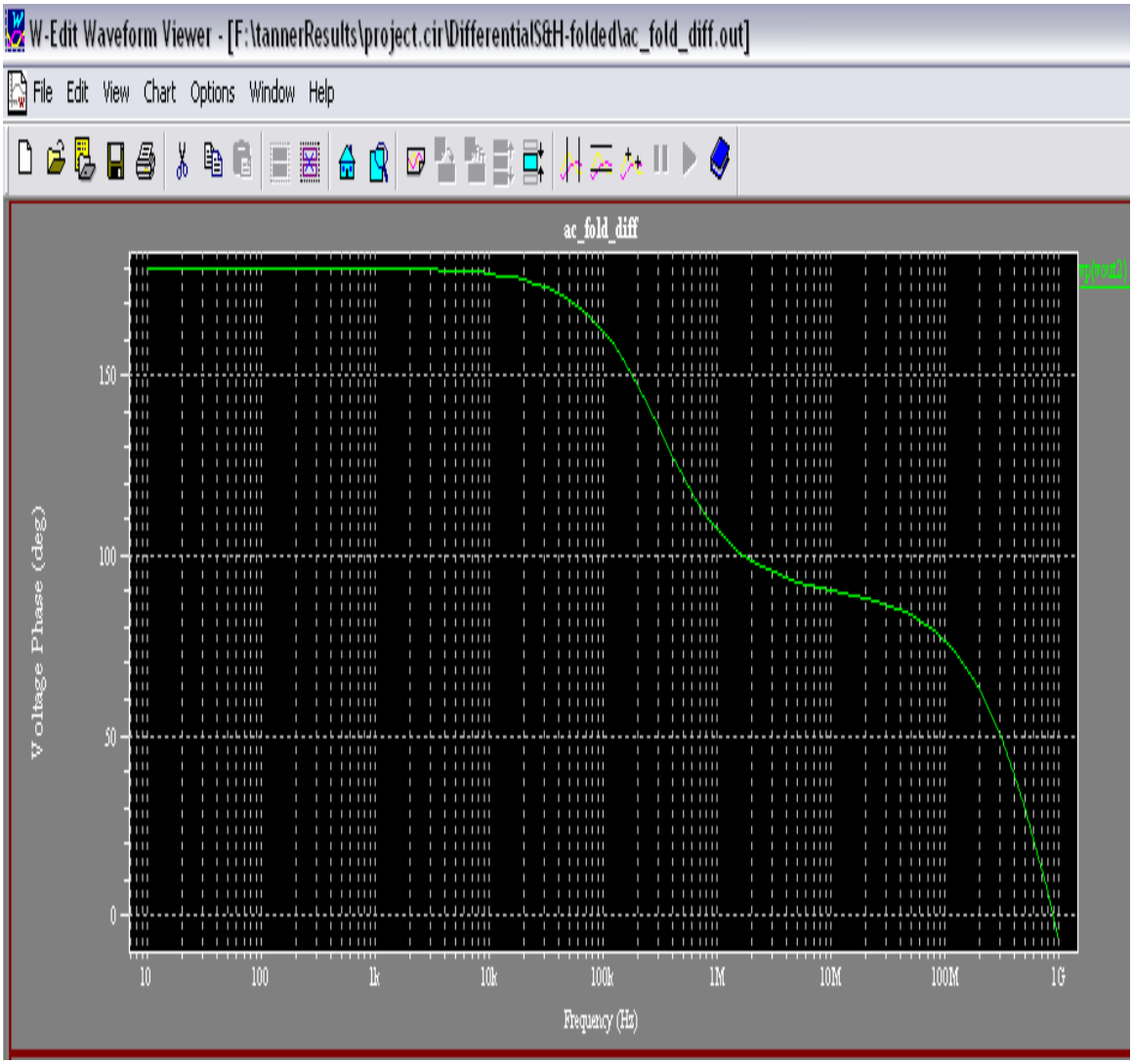


Figure 5.11: Phase Response of the Fully Differential Folded Cascode Op-amp

Our final step towards the end of results will be to prove the accuracy of designed sample and hold circuit, which infact is the most important result because it decides how the system will behave in the noisy environment.

In order the to prove the accuracy we will use the well known SNR formula for a N-bit ADC which was also mention in the first chapter and repeated here for convenience.
SNR actually is the figure of merit for this circuit.

For a N-bit ADC the SNR should be given by below expression:

$$SNR = 6.02N + 1.76 \text{ dB}$$

Approximately we can write the above expression as

$$SNR = 6N$$

The SNR of the designed sample and hold circuit can only be calculated from the Fast Fourier Transform (FFT) of the output sample signal. Here, we have calculated 32-point FFT using the table below.

Table 5.2 32-points for FFT calculation.

1	Time	Vout(1)	Vout(2)	Vout(1)-Vout(2)
2				
3	4.44E-08	2.44E+00	2.3536	-0.0899
4	9.09E-08	2.55E+00	2.2632	-0.2834
5	1.42E-07	2.64E+00	2.1734	-0.4642
6	1.93E-07	2.72E+00	2.0931	-0.6276
7	2.46E-07	2.80E+00	2.0335	-0.7766
8	2.95E-07	2.85E+00	1.98E+00	-8.76E-01
9	3.50E-07	2.90E+00	1.9487	-9.51E-01
10	3.96E-07	2.92E+00	1.9265	-9.91E-01
11	4.45E-07	2.91E+00	1.9218	-9.92E-01
12	4.98E-07	2.90E+00	1.95E+00	-9.54E-01
13	5.46E-07	2.86E+00	1.98E+00	-8.81E-01
14	5.91E-07	2.78E+00	2.0093	-7.74E-01
15	6.40E-07	2.72E+00	2.0802	-6.37E-01

16	6.90E-07	2.64E+00	2.1646	-4.76E-01
17	7.39E-07	2.55E+00	2.2564	-2.96E-01
18	7.90E-07	2.46E+00	2.3528	-1.05E-01
19	8.40E-07	2.36E+00	2.45E+00	9.09E-02
20	8.92E-07	2.26E+00	2.5471	2.83E-01
21	9.42E-07	2.17E+00	2.6382	4.64E-01
22	9.92E-07	2.09E+00	2.719	6.27E-01
23	1.04E-06	2.01E+00	2.7807	7.67E-01
24	1.10E-06	1.99E+00	2.8476	8.54E-01
25	1.15E-06	1.96E+00	2.887	9.28E-01
26	1.20E-06	2.05E+00	2.8121	9.92E-01
27	1.25E-06	1.93E+00	2.9214	9.94E-01
28	1.30E-06	1.94E+00	2.9019	9.57E-01
29	1.35E-06	1.98E+00	2.86E+00	8.84E-01
30	1.40E-06	2.03E+00	2.8064	7.76E-01
31	1.44E-06	2.08E+00	2.719	6.39E-01
32	1.50E-06	2.18E+00	2.6527	4.76E-01
33	1.55E-06	2.27E+00	2.5621	2.96E-01
34	1.58E-06	2.37E+00	2.48E+00	1.12E-01

These 32-points are being derived from the output sampled signal. Each sample is the value of the settled sampled signal. In the figure5.12 there are 32 samples in one period. These 32 points can be obtained by using the below given formula

$$f_{signal} = \frac{f_{sampling}}{32}$$

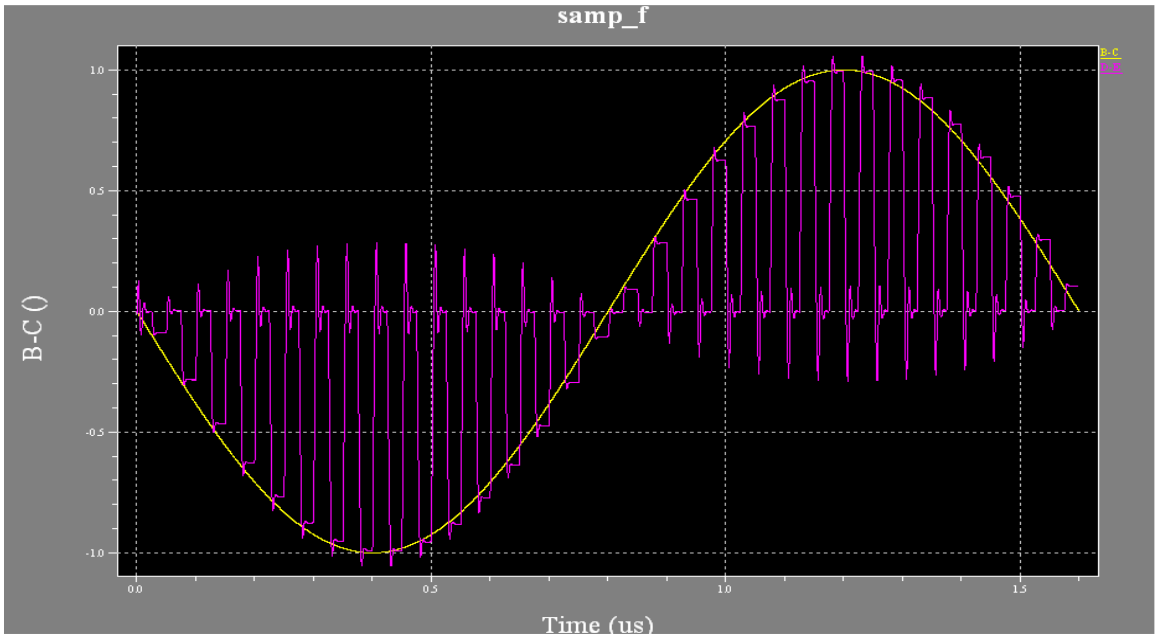


Figure 5.12: 32-bit sampled output waveform

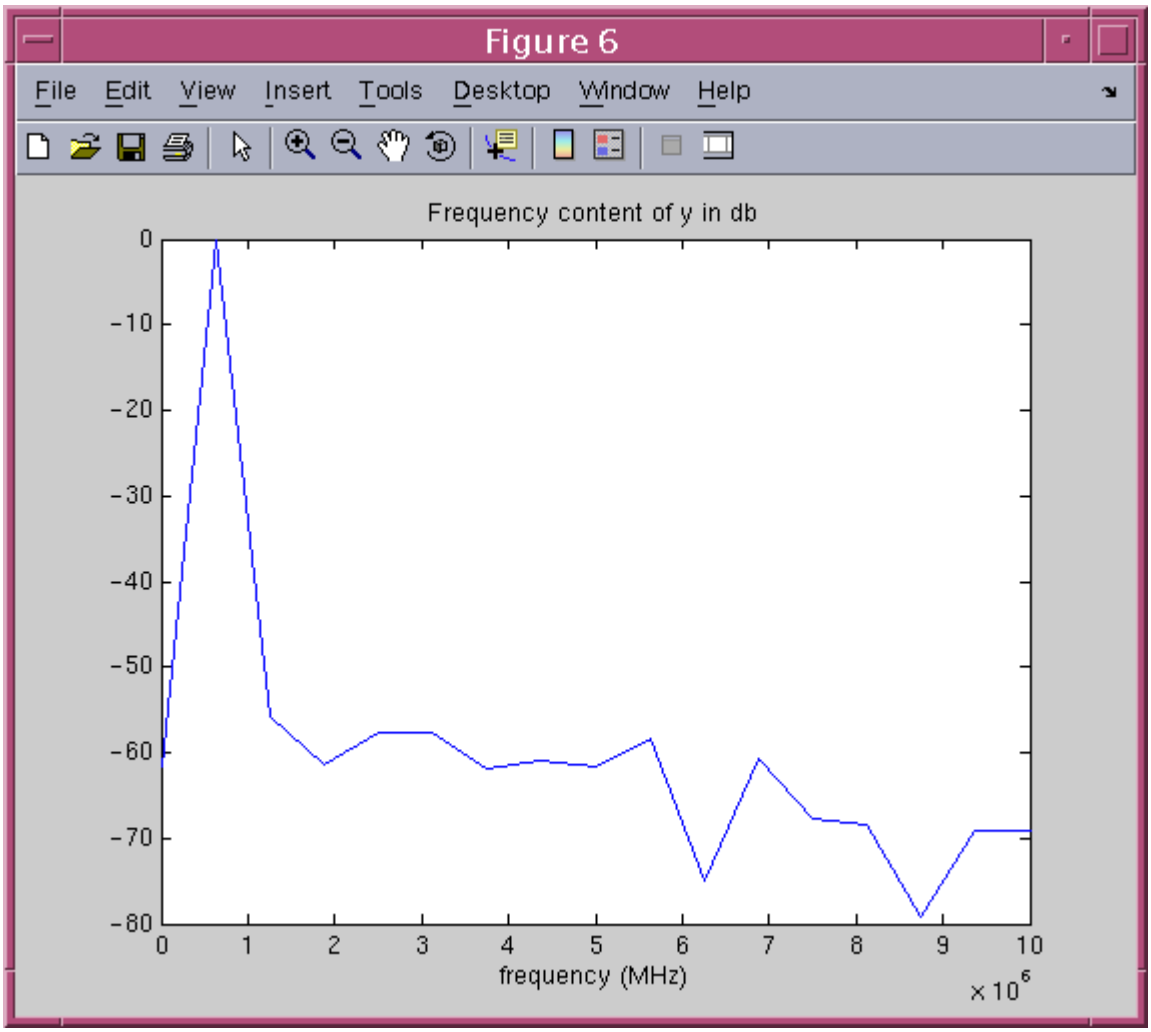


Figure 5.13: Fast Fourier Transform (FFT) of the output sampled signal.

The above figure shows the 32-point FFT of the output sampled signal. The graph is being normalized with respect to the fundamental frequency lobe. Here the main observation is the higher lobe which represents the signal frequency and is surrounded by noise present in the system between -50db to 60db. The SNR of the noise signal with respect to fundamental frequency lobe will give the accuracy obtained using equation described above which will be used as a result to verify the accuracy desired for the system.

The SNR of the system is found to be 57db.

Now again using the equation SNR=6N, where N is the bit accuracy of the system. For our design, N=10. Using this formula we can get the N of the actual designed system. Calculating, we get N=9.5. So this is fairly accurate design and hence the proof.

Table below shows is the specification achieved for the designed system.

Table 5.3: Achieved Specifications of Sample and Hold Circuit

Supply Voltage	5V
Sampling Frequency	20MHz
Input Frequency	2MHz
Power	1.8mW
Output Range/Swing	-1.5V to +1.5V
Gain of Op-amp	68dB
Bandwidth	240MHz
Phase Margin	57degrees
Slew Rate	130V/uA
Accuracy	9.5 bits

Chapter 6

Conclusion and Future Scope

6.1 Conclusion

With the increasing demand for high-resolution and high-speed in date acquisition systems, the performance of the S/H circuits is becoming more and more important. This is especially true in ADCs since the performance of S/H circuits greatly affects the speed and accuracy of ADCs. The fastest S/H circuits operate in open loop, but when such circuits are implemented in CMOS technology, their accuracy is low. S/H circuits that operate in closed loop configuration can achieve high resolution, but their requirements for high gain circuit block, such as an op-amp, limits the speed of the circuits [8]. As a result, better and faster S/H circuits must be developed.

The simulation of sample hold circuit design provided an insight into the problems that occurs for an analog circuit. The values obtained by hand calculations had to be modified a bit during simulation phase in order to meet the specification. Also, we can imagine the amount of complexity that will arise when a analog circuit is actually implemented. This project however bridged the gap between the theory and the practical design to considerable extent.

6.2 Future Scope

Sample-and-hold (S/H) is an important analog building block that has many applications. The simplest S/H circuit can be constructed using only one MOS transistor and one hold capacitor. However, due to the limitations of the MOS transistor switches, errors due to charge injection and clock feedthrough restrict the performance of S/H circuits. As a result, different S/H techniques and architectures are developed with the intention to reduce or eliminate these errors.

At the same time, the employment of low-voltage in VLSI technology requires that the analog circuits be low-voltage as well. As a result of this, new researches in analog circuits are now shifted from voltage-mode to current-mode. The advantages of current-

mode circuits include low-voltage, low-power, and high-speed [23]. Therefore, future researches of S/H circuit should also shift toward current-mode S/H techniques.

But still more new S/H techniques and architectures need to be proposed in order to meet the increasing demand for high-speed, low-power, and low voltage S/H circuits for data acquisition systems.

Reference

[1] A. V. Oppenheim and R. W. Schafer, Discrete-Time Signal Processing. Prentice Hall Signal Processing Series. Prentice Hall, Englewood Cliffs, NJ, USA, 1989.

[2] B. Razavi, “Design of Sample-and-Hold Amplifiers for High-Speed Low-Voltage A/D Converters”, Proceedings of the IEEE Custom Integrated Circuits Conference, pp. 59-66, 1997.

[3] B. Razavi, “Design of a 100-MHz 10-mW 3V Sample-and-Hold Amplifier in Digital Bipolar Technology", IEEE Journal of Solid-State Circuits, vol. 30, pp. 724-730, July 1995.

[4] B.Razavi, “Design of Analog CMOS Integrated circuits”, Tata McGraw-Hill Publication, 2002

[5] B. K. Ahuja, “An Improved Frequency Compensation Technique for CMOS Operational Amplifiers,” IEEE Journal of Solid-State Circuits, vol. 18, pp. 629-633, Dec. 1983

[6] C. Eichenberger and W. Guggenbuhl, “Dummy Transistor Compensation of Analog MOS Switches”,IEEE Journal of Solid-State Circuits, vol. 24, pp. 1143-1146, Aug. 1989.

[7] C. Moreland et al., “A 14b 100MSample/s 3-Stage A/D Converter", ISSCC Digest of Technical Papers, pp 34-35, Feb. 2000

[8] C. R. Ryan, “Applications of a Four-Quadrant Multiplier,” IEEE J. Solid-state Circuits, vol. SC-5, pp. 45-48, Feb. 1970.

[9] D. Kelly et al., \A 3V 340mW 75MSPS CMOS ADC with 85dB SFDR at Nyquist", ISSCC Digest of Technical Papers, pp 134-135, Feb. 2001

[10] Eric Vittoz et al., “Charge Injection in Analog MOS Switches", IEEE Journal of Solid-State Circuits, vol. sc-22, pp. 191-1097, Dec. 1987.

[11] H. Pan et al., “A 3.3V, 12b, 50MSample/s A/D Converter in 0.6um CMOS with over 80 dB SFDR", ISSCC Digest of Technical Papers, pp 40-41, Feb. 2000.

[12] L. Dai, R. Harjani, “CMOS Switched-Op-Amp-Based Sample-and-Hold Circuit IEEE Journal of Solid-State Circuits, vol. 35, no. 1, pp. 109-113, January 2000.

[13] L. Singer et al., \A 12b 65MSample/s CMOS ADC with 82 dB SFDR at 120MHz", ISSCC Digest of Technical Papers, pp 38-39, Feb. 2000.

[14] M. A. M. Zin, et al., .A High-Speed CMOS Track/Hold Circuit, IEEE International Conference on Electronics, Circuits and Systems 1999, vol. 3, pp. 1709-1712, 1999.

[15] M. Shingagawa et al., "Jitter Analysis of High-Speed Sampling Systems", IEEE Journal of Solid State Circuits, vol. 25, pp. 220-224, Feb. 1990.

[16] M. Waltari and K. Halonen, "Timing Skew Insensitive Switching for Double Sampled Circuits", IEEE International Symposium on Circuits and Systems, May 1999.

[17] P. Petschacher et al, "A 10-b 75-MSPS Subranging AD Converter with Integrated Sample and Hold," IEEE J. Solid-State Circuits, vol. SC-25, pp. 1339-1346, Dec. 1990.

[18] P. Real and D. Mercer, "A 14b Linear, 250nsec Sample and Hold Subsystem with Self-Calibration," ISSCC Dig. Tech. Papers, pp. 164-165, Feb. 1991.

[19] P. R. Gray and R. G. Meyer, "MOS Operational Amplifier Design: A Tutorial Overview," IEEE Journal of Solid-State Circuits, vol. 17, pp. 969-982, Dec. 1982.

[20] P. R. Gray and R. G. Meyer, "Analysis and Design of Analog Integrated Circuits, 4th ed, John Wiley & Sons, Inc, 2005.

[21] R. J. Baker, CMOS: Circuit Design, Layout and Simulation, 1st ed. Boise, ID: Wiley-IEEE, 2005.

[22] R. J. Baker, "Design of High-Speed CMOS Op-Amps for Signal Processing,"IEEE/EDS WMED, April, 2005.

[23] Y. Sugimoto, "A 1.5-V Current-Mode CMOS Sample-and-Hold IC with 57-dB S/N at 20 MS/s and 54-dB S/N at 30 MS/s," IEEE Journal of Solid-State Circuits, vol. 36, no.4, pp. 696-700, April 2001.

[24] 14-Bit, 40 MSPS/65 MSPS A/D Converter. AD6644 Data Sheet, Rev. 0. Analog Devices, Inc. May 2000.

[25] 14-Bit, 80 MSPS Analog-to Digital Converter. AD6645 Preliminary Datasheet, Rev. PRb. Analog Devices, Inc., July 2001.

[26] 14-Bit, 105 MSPS A/D Converter. AD6645 Preliminary Datasheet, Rev. PRa. Analog Devices, Inc., Sept. 2001.