UNIVERSITY OF CALIFORNIA

Los Angeles

Mixed-Signal Circuit Design Driven by Analysis: ADCs, Comparators, and PLLs

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Electrical Engineering

by

Hao Xu

2018

ABSTRACT OF THE DISSERTATION

Mixed-Signal Circuit Design Driven by Analysis: ADCs, Comparators, and PLLs

by

Hao Xu
Doctor of Philosophy in Electrical Engineering
University of California, Los Angeles, 2018
Professor Asad Abidi, Chair

Mixed signal circuit design often involves circuits that are time-varying or highly non-linear, which further results in systems that are difficult to characterize using established methodologies for linear time-invariant systems, thus designers are more than often forced to rely on intensive simulations for design. This dissertation explores design optimization for comparators, phase locked loops and ADC from three different perspectives.

First, a complete analysis for regenerative comparators is presented including noise, offsets and speed for the first time. Despite the fact that the comparators are time-varying and regenerative with infinite gain, simple equivalent circuits still accurately capture their operation. Design guideline are provided for different comparator architectures.

Second, a linearized analysis for phase locked loops using bang-bang phase detectors is presented. The high non-linear bang-bang phase detector is ascribed to an effective gain, whose physical meaning is interpreted in signal space. Closed form expressions for loop gain, output jitter and phase noise profile are obtained using transfer functions for the first time. Design guidelines are also provided.

Last, a 2.5GS/s 10bit 65mW ADC in 28nm CMOS FD-SOI without active amplifier and intensive digital calibration is presented. This highlights the potential of circuit design based on complete understandings. The fabricated ADC with considerably less complexity achieves comparable performance with state-of-arts. Different imperfections are quantitatively studied and compared with measurement.

The dissertation of Hao Xu is approved.

Danijela Cabric

Sudhakar Pamarti

Ken Yang

Asad Abidi, Committee Chair

University of California, Los Angeles

2018

To my parents.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

It had not been my original intention to pursue a Ph.D when I started graduate study at UCLA. I probably would not have chosen to do so if I had known that it would take over eight years. Although this journey has been quite rocky and turns out to be longer than I had expected, to me, it is still worthwhile.

I have been very fortunate to be advised by professor Abidi. Beyond his well known magical power of simplifying seemingly complicated circuits, what I learned the most from him is his dedication of pursuing knowledge and his respect to engineering. There are many reasons we engineers are tied to engineering, but ultimately, we do it because we love it. This keeps resonating with me. Not many people still possess such pure passion nowadays. I would also like to thank professor Yang, professor Pamarti and professor Cabric for kindly serving in my committee and offering many feedbacks.

Through my study, Hui Pan from Broadcom has helped me not only with his expertise in circuit design, but also in personal life and career choice. I am greatly thankful to him. Glenn Chang and Tony Wang from Maxlinear went extra miles to arrange lab testing for my chip at a critical time when sponsorship to academia was scarce. This help will be always appreciated.

Along my time in industry, I have had the luxury to work with many talented designers who made impact to my study and career. I spent 2011 summer at Tensorcom, where I met Jiashu Chen, whose knowledge and determination still impress me today. During 2012 and 2013 summers, I had the chance to work on a bit of RF design under the guidance of Hooman Darabi and David Murphy at Broadcom. Although not directly related to my Ph.D topic, I learned a great deal from them. During the summer of 2014, I interned under the guidance of Ray Gomez and Hui Pan at Broadcom, which was the first time I was exposed to practical ADC design. In 2017 summer, I worked at Mediatek Irvine office on high speed ADCs with Tamer Ali, Henry Park, E-Hung Chen and Ramy Awad. Since 2017 fall, I have become a full-time employee at Broadcom PLP group. I greatly appreciate the help from my supervisors Delong Cui and Jun Cao, and many colleagues who carried on the project

while I was busy with my graduation.

At UCLA I came across many friends. I had a good time playing basketball with Ning-Yi Wang, Hao Wu, Zhao Yang, Yafeng Zhang, Fengbo Ren and other friends, though later on this activity was more or less traded for study. Professor Tetsuya Iizuka from the University of Tokyo developed the analysis of passive T/H during his stay in our group, which was later adopted in the prototype ADC design. His diligence has been great inspiration to me. I also had many fruitful discussions with my fellow group members, Wenlong Jiang, Alvin Chen, Qaiser Nehal, Alireza Yousefi, Jiacheng Pan, Dihang Yang, Weiyu Leng, Mohammad Fayazi and Kejian Shi. Especially, Dihang Yang helped on the layout of the ADC chip and Wenlong Jiang offered significant help in the final verification. Xiao Wang, my roomate for two years, has created quite some hilarious and memorable moments. The section would become overwhelmingly long to include names of all my friends over the world who have supported me. These support have been and will always be appreciated.

My parents have offered me their all and provided me opportunities that they have never had. I could not ask for any more. This dissertation is dedicated to them.

My final thanks are for my wife Wenlu Jiang. Without exaggeration, she brought light to me in the darkest days.

VITA

2006–2010 B.S. Microelectronics
Fudan University

2010–2012 M.S. Electrical Engineering
University of California, Los Angeles

2012–2018 Ph.D. Electrical Engineering
University of California, Los Angeles

PUBLICATIONS

H. Xu and A. Abidi, 'Design Methodology for Phase-Locked Loops using Binary (Bang-Bang) Phase Detectors', *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 7, pp. 1637-1650, July 2017.

A. Abidi and H. Xu, 'Understanding the Regenerative Comparator Circuit,' *Proceedings of the IEEE 2014 Custom Integrated Circuits Conference*, San Jose, CA, 2014, pp. 1-8.

D. Murphy, H. Darabi and H. Xu, 'A Noise-Cancelling Receiver Resilient to Large Harmonic Blockers,' *IEEE Journal of Solid-State Circuits*, vol. 50, no. 6, pp. 1336-1350, June 2015.

D. Murphy, H. Darabi and H. Xu, '3.6 A Noise-Cancelling Receiver with Enhanced Resilience to Harmonic Blockers,' *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, San Francisco, CA, 2014, pp. 68-69.

D. Murphy, H. Darabi and H. Xu. 'Generating high-frequency, non-overlapping clocks.' U.S. Patent No. 8,618,859. 31 Dec. 2013.

CHAPTER 1

Introduction

Mixed-signal circuit design, including ADCs and PLLs for example, is arguably the field that has the most 'unknowns'. Since it is where analog circuits and digital circuits intersect, it often involves circuits that are time-varying with hard non-linearity, resulting in barriers for systematic designs as circuit designers are usually trained to deal with linear time-invariant(LTI) systems. The design space might seem overwhelmingly broad because of lack of theoretical tools. The motivation of this dissertation is to tackle the exact problems related to the time-varying nature and hard non-linearity in mixed-signal circuit design. The goal is to develop certain analytical frameworks so that designing these circuits are no longer iterations of trial and error, but rather linear processes where directions of optimization are well defined. Such frames have been built for regenerative comparators and the bang-bang PLLs, both of which are time-varying and highly non-linear. The study on comparators presented in this dissertation is the first complete study that answers all questions related to comparators. The study on bang-bang PLLs in this dissertation is the first design-oriented analysis that provides straightforward expressions and guidelines. The approach of design by understanding is finally applied to a prototype ADC design where design trade-offs are all quantitatively explored.

Chapter two studies regenerative comparators. Despite the hard non-linearity and time-varying nature of these circuits, simple equivalent circuits have been developed to accurately characterize offsets, noise and speed. The long standing question related to dynamic offsets, caused by common mode to differential mode conversion due to circuit mismatches, is visualized using phase plane plots and explained with simple circuits. Different comparator architectures are quantitatively compared and design guidelines are provided.

Chapter three presents a frequency domain analysis for phase locked loops using bang-bang phase detectors. The hard nonlinearity of bang-bang phase detectors complicates loop design as unlike linear phase detector that has a well defined gain, the concept of linear gain cannot be readily ascribed to bang-bang phase detectors. Previous published works either apply time domain analysis using difference equations or rely on numerical iterative solvers to analyze the loop behavior, neither of which offers straightforward design insights. A linearized analysis has been proposed in this dissertation that is complete and self-consistent. It enables the manual design of frequency synthesis PLLs for loop bandwidth, output phase noise and minimum jitter.

Chapter four demonstrates a 2.5GS/s 10bit 65mW ADC design without active amplifier and intensive calibration. While most state-of-arts with similar sampling frequency and resolution apply rather sophisticated digital calibration, this design is driven by complete circuit understandings and involves minimum calibration. It is shown that a fully optimized design, with considerably less design complexity and design efforts, is able to achieve comparable performance with designs with more intensive calibration. Various imperfections are quantitatively studied and characterized. Digital calibration is a powerful tool to improve performance of optimized designs when analog circuits hit certain limits, but is not a universal solution for any design that is not fully optimized. The analytical frame presented in this dissertation enables designers to make design choices early on.

Chapter five summarizes the key contributions and proposes future improvements.

CHAPTER 2

Understanding Regenerative Comparators

The regenerative comparator circuit which lies at the heart of A/D conversion, slicer circuits, and memory sensing, is unstable, time-varying, nonlinear, and with multiple equilibria. That does not mean, as this paper shows, that it cannot be understood with simple equivalent circuits that reveal its dynamics completely, and enable it to be designed to specifications on static and dynamic offset and noise. The analysis is applied to the StrongArm latch.

2.1 Introduction

Flip-flops used as regenerative amplifiers are found everywhere in electronic circuits. It is well-understood that they have two stable states, and that given sufficient time, the circuit, which involves positive feedback, will regenerate an input voltage unbalance to reach one of these states.

The latched comparator must be symmetric by its very nature, since its binary states are symmetrical. Practical unbalances in the circuit arising from transistor and load mismatch lead to uncertainty in the regenerated binary output when a small analog input is applied. This problem of offset is well-known to circuit designers, and since the observed offsets can be much larger than in simple linear amplifiers, a certain mystery attends to "dynamic offsets" that appear only in a latched comparator [12, Sec. 4.14].

Despite vast amount of publications, offsets, especially dynamic offsets, are not well understood. The underlying difficulties come from two aspects. First, offsets arise from circuit imbalances, which disturb the symmetry of circuits. This incurs significant challenge to characterizing circuit behavior. Based on 'half circuit theory', a fully symmetric circuit,

whether linear or non-linear, can be decomposed into the 'common mode' half and the 'differential mode' half circuits. These two modes are independent of each other. This decomposition usually reduces the number of variables involved in calculation and greatly simplifies circuit analysis. However, this decomposition cannot be directly applied to circuit with imbalances. Second, in almost all cases, latched comparators are time-varying circuits that experience large voltage or current excursions. This further increases the complexity of calculation and analysis.

Since circuit with imbalances were considered not applicable for 'half circuit' decompositions, solving full-blown differential equation including all state variables seemed to the only viable choice. Although this approach renders correct results, with overwhelmingly complicated equations, they fall short of offering design insights. The complexity involved in equation solving has further restricted this approach to only static latches [13, 14, 15], which can be represented with time invariant equivalent circuits. But most regenerative comparators are time-varying circuits with multiple phases. A single equivalent circuit is insufficient to capture all circuit operations. With a series of multiple equivalent circuits corresponding to each phase, the differential equations quickly becomes too mathematically intensive to provide any insight.

StrongARM latch, first proposed by [16], is one of the most widely used regenerative comparators. Because of the combined challenges from circuit imbalances and time-varying circuit operations, it is still not well understood. [17, Eq. (7)] treats a strongARM latch the same as a linear differential amplifier and thus fails to distinguish between static and dynamic offsets. [18] only considers a certain window of the strongARM operation and ignores its time-varying nature. The time-varying nature of regenerative comparators was firstly appreciated in publications investigating noises [19, 20]. [21] expands the frame to include dynamic offsets caused by capacitor loading imbalances, yet the results are only partially correct because of lack of theoretical tool dealing with circuit imbalances.

In this paper we present a systematic design-oriented analysis for regenerative comparators. This dissertation is based on our previous work [22] and we contribute the following beyond already published work:

1. Realize the common-mode to differential-mode coupling in an asymmetric circuit. This enables us to convert an asymmetric circuit to a symmetric circuit with extra common-mode to differential-mode disturbance. And then the well established half circuit analysis can be applied. Simple and accurate results are then obtained.
2. Phase plane is used to visualize regeneration. By using one trajectory as a frame of reference, circuit unbalances can in most cases be modelled by a sequence of linear, time-varying equivalent circuits that capture, piecewise, the variation of the circuit with time.
3. Specify the phases involved in a strongARM latch and quantify the preamplification brought by its circuit operation. Along with common-mode to differential-mode coupling, this enables us to calculate and understand both static and dynamic offsets without overcomplicated mathematical derivations.
4. Extend the analysis to noises in strongARM latch. We present a simple analysis using the concept of equivalent noise bandwidth (NBW).
5. The analysis leads to design guidelines for strongARM latches including offset, noise and speed. We quantitatively present the benefits of correctly choosing common-mode and the degradations when it is not well designed.
6. Analysis is verified against both simulations and measurements.

Regenerative comparators have various forms of configurations. StrongARM is chosen for analysis in this dissertation because of two reasons. First, it has a delicate circuit operation with substantial embedded preamplification, which further suppresses noise and offset. Second, because of its delicate configuration, it is arguably more challenging to characterize. Although in this dissertation we focus our discussion to static latches and strongARM latches due to space limit, the adopted theoretical approach extends to all other comparator architectures, including the nowadays popular two-stage or double tail comparators [23, 24, 25, 26].

2.2 The Static Latch

Any discussion of a regenerative amplifier must start with the CMOS static latch (Fig. 2.1). This is a classic circuit: simple, uncluttered, and therefore easily understood. The only choice

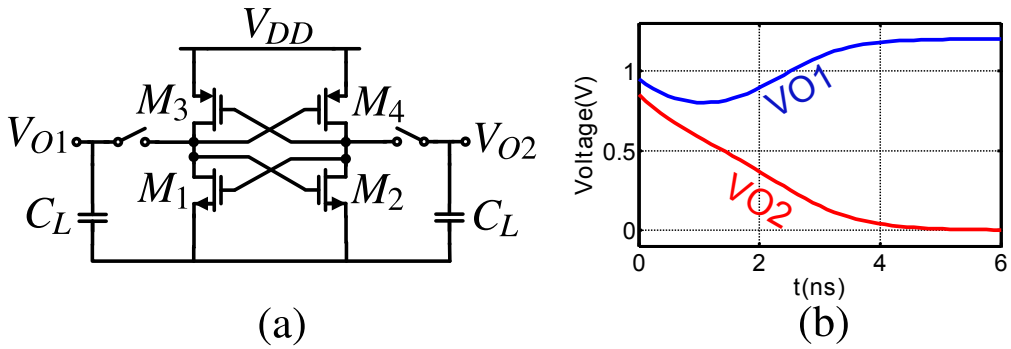


Figure 2.1: (a) Schematic of a static latch; (b) Voltage waveforms in a static latch.

lies in the method whereby a small analog input is coupled into the latch without disturbing regeneration. Here, the two load capacitors C_L are pre-charged to a common voltage, on which is superposed a small differential voltage. That is, the voltage V_{O1} is slightly larger than V_{O2}, while their average voltage, in this example, is chosen slightly lower than V_{DD}. The pre-charged capacitors are switched into the latch (it will be seen that there are *only* two nodes in the circuit), and the circuit regenerates this difference to the voltage rails. The regeneration waveform (Fig. 2.1) is familiar to almost everyone who designs circuits.

The existing literature does not give a satisfactory explanation for the one striking feature of every regenerative waveform: Why do V_{O1} and V_{O2} fall together, and then at some critical time, they peel away and separate? To find the answer we create a phase portrait of this circuit [27, Ch. 11]. The independent capacitors define two state variables, but using the capacitor voltages as state variables can obscure the essential properties. Instead, we take the average of the capacitor voltages as one state variable, the common-mode voltage V_{OC}, and their difference V_{OD} as the other. These are orthogonal quantities, in the sense that one can change while the other remains constant. The phase plane has been used before to investigate latch dynamics [28, 29, 30], but not defined by the circuit modes. As will soon be clear, this choice makes all the difference.

2.2.1 Phase Plane

The phase plane is defined by axes V_{OC} vs. V_{OD} (Fig. 2.3). It is covered by a *vector field* that is signified at every point (V_{OD}, V_{OC}) on the plane by an arrow with the magnitude and direction of the ratio $(dV_{OC}/dt) \div (dV_{OD}/dt)$, where each time derivative at that point is obtained from the circuit equations. Starting from any initial condition in the plane, there develops by connecting the vector field a unique integral curve which depicts graphically how the state variables will change with time.

The vector field that fills the phase plane can equally well describe a nonlinear differential equation, or a linear one. Sometimes linearity will apply in a limited region of the plane. In that region any integral curve can be decomposed into a superposition of two eigenvectors. In turn, each eigenvector can be associated with the natural response of some linear circuit. For the static latch, as it turns out, linearity is a very good assumption over much of the excursion of V_{O1}, V_{O2}, particularly over the *one* integral curve that will interest us. If $\beta_N = \beta_P$ where $\beta \triangleq \mu C'_{ox} W/L$ for the CMOS inverters, then the overall G_m of the inverter remains almost constant except near $\frac{1}{2}V_{DD}$. At low V_{DD} this appears a small local deviation. The equivalent circuit that is produced by replacing each CMOS inverter with a constant G_m voltage-controlled current source Fig. 2.2 resembles a differential amplifier, but the cross-coupled controlling voltages identify it as a flip-flop. C_p is a net cross-coupling capacitor between the two nodes arising from FET capacitance. We have deliberately left out the output conductance g_{ds} of the FETs since it makes little difference to the analysis that now follows.

2.2.2 Equivalent Circuits for Modes

A decomposition into common-mode and differential mode will also guide the search for meaningful equivalent circuits. Since these modes are independent, any response of this linear circuit can be decomposed into a superposition of the two modes. If the symmetric equivalent circuit (Fig. 2.2) is operating purely in common mode, then $V_{O1}(t) = V_{O2}(t)$. Since by symmetry no currents will flow through C_p, the circuit can be bisected into two

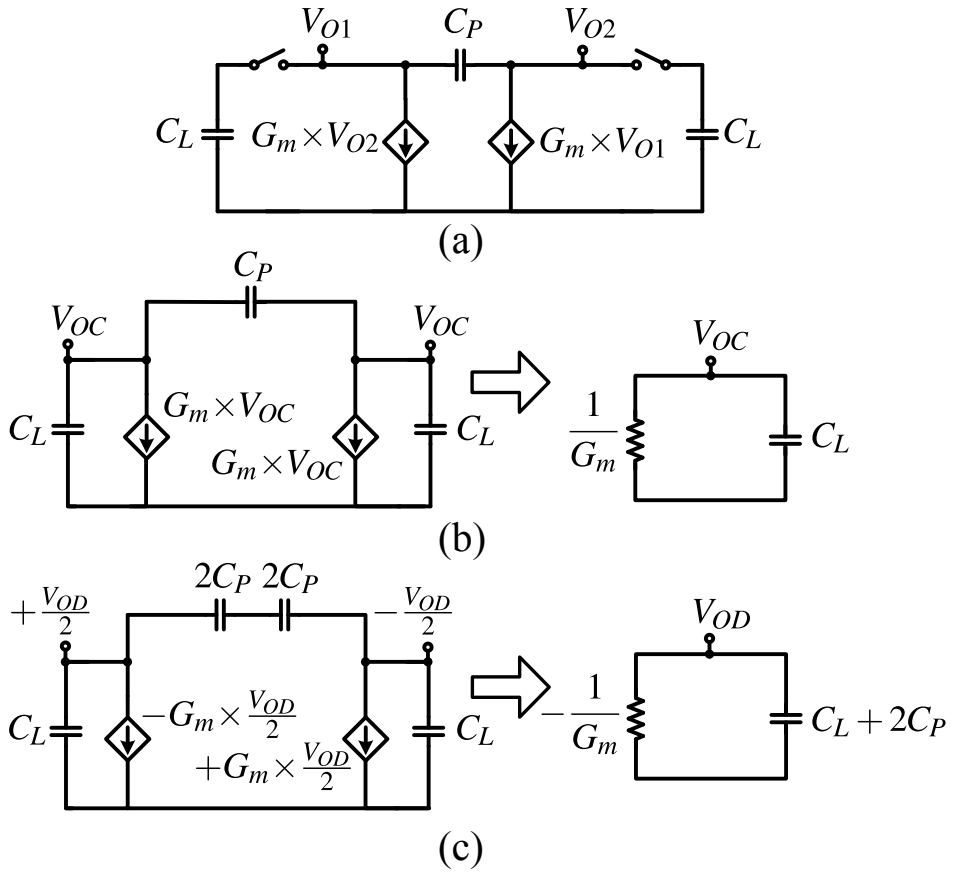


Figure 2.2: (a) Equivalent circuit of a static latch; (b) Common-mode equivalent circuit; (c) Differential mode equivalent circuit.

half circuits. It is clear that the natural response of the two halves of the circuit in common mode is stable, with a pole located at

$$s_c = -G_m/C_L. \tag{2.1}$$

On the other hand, if the symmetric circuit is operating purely in differential mode, then $V_{O1}(t) = -V_{O2}(t)$. Symmetry now dictates that the mid-plate potential in the capacitor C_P remains zero. After bisection into half circuits, each half consists of a capacitance $C_L + 2C_P$ across a negative conductance $-G_m$. Now the natural response of the two halves is *unstable*, defined by a pole located at

$$s_d = +G_m/(C_L + 2C_P). \tag{2.2}$$

Each mode will be stimulated by its own initial condition. Therefore the time response to any initial condition $V(0) = V_{OC}(0) + V_{OD}(0)$ may be expressed as a superposition of the

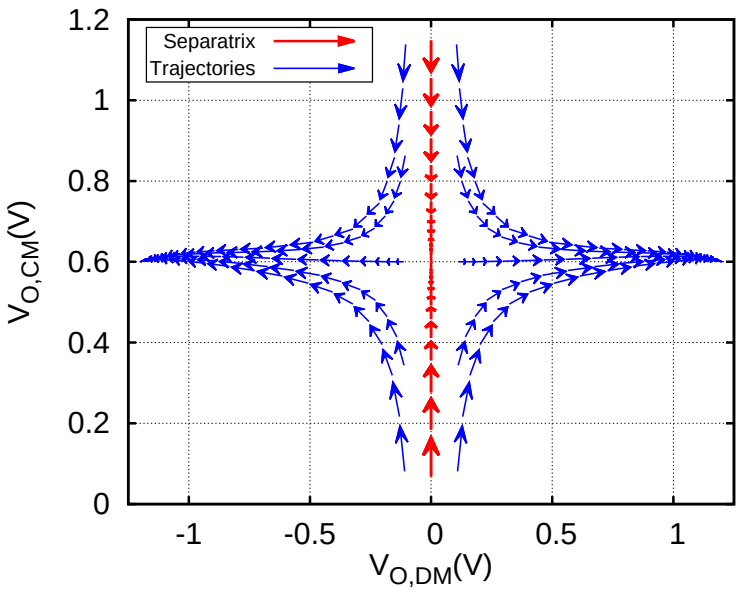


Figure 2.3: Phase plane plot of a static latch. V_{DD}=1.1 V.

common and differential modes that are stimulated:

$$V_{O1,2}(t) = V_{OC}(0)\exp\left(-\frac{G_m t}{C_L}\right) \pm \frac{1}{2}V_{OD}(0)\exp\left(+\frac{G_m t}{C_L + 2C_P}\right) \tag{2.3}$$

This expression lends understanding to the example integral curves plotted on the phase plane. The vertical trajectories start from an initial condition $V_{O1} = V_{O2}$ that does not stimulate the differential mode. The trajectories lead into an equilibrium point at the coordinates $(0, \frac{1}{2}V_{DD})$. On the other hand if the output nodes are initialized so that $V_{OD} \neq 0$ but $V_{OC} = \frac{1}{2}V_{DD}$, that is, the circuit is released from an initial condition that is in purely differential mode, the horizontal trajectory in the phase plane accelerates away from this equilibrium point. If the circuit were truly linear this trajectory head towards $V_{OD} \to \pm\infty$. But in the actual circuit the node voltages cannot exceed the supply or ground, so as these trajectories approach the coordinates $(+V_{DD}, \frac{1}{2}V_{DD})$ and $(-V_{DD}, \frac{1}{2}V_{DD})$ they slow down to come to rest at one of these two equilibrium points. These last two equilibria are *stable*. However, the equilibrium at $(0, \frac{1}{2}V_{DD})$ is *metastable*: only if the circuit is initialized along the common-mode axis will it approach the equilibrium; otherwise, for any other initial condition, it will be deflected away from it towards one of the two stable equilibria. In phase plane terminology, the metastable equilibrium defines a saddle point.

Using the phase plane, we are able to understand the characteristic time-domain waveforms of a static latch. In a typical use, the latch is initialized with both nodes connected to, or biased close in voltage to, the power supply. This defines a large initial common mode. A small differential voltage is superimposed to direct the latch regeneration. On the phase plane this may correspond to the initial condition of the trajectory on the upper right. The large initial common mode will decay towards the metastable point, but the small differential mode will grow exponentially. The resulting trajectory is a superposition of the stable eigenvector which lies on the vertical axis, and the unstable eigenvector on the horizontal axis. Because of the large initial condition, the stable eigenvector dominates at first causing V_{O1} and V_{O2} to decay together. Then, as the unstable eigenvector grows, the differential voltage becomes dominant, causing the two voltages to split apart until the circuit reaches a stable equilibrium point. The turnaround point in the $V_{O1}(t)$ waveform, defined by its minimum value, corresponds to that point on the phase plane trajectory where the slope of the vector field is 0.5 in magnitude.

2.2.3 Circuit Imbalances

Offsets are important when the latched comparator is used in analog-to-digital conversion. Indeed, even in memory sense applications, offsets in the sense amplifier can be so large as to pose a threat to reliable readout. In a nominally symmetric circuit, offsets arise from parameter mismatch in corresponding pairs of elements, such as in the threshold voltage V_t of the NMOS pair, or in the capacitance of the loads C_L. All unbalances appear as an input-referred offset voltage, which we now seek to estimate by simple analysis.

The effects of unbalances in the comparator are most neatly analyzed with half circuits. Middlebrook shows how to treat parameter unbalances in *static* symmetric circuits [31] as equivalent half circuits; we generalize this to *dynamic* circuits.

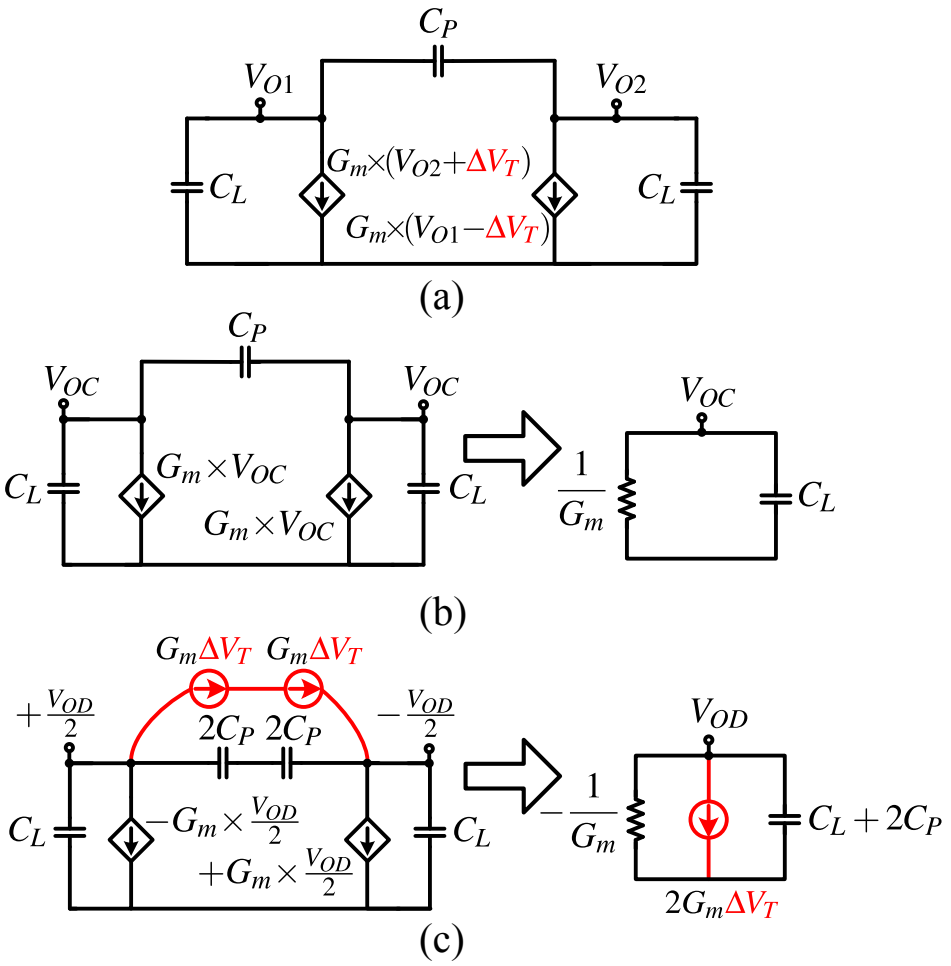


Figure 2.4: (a) Equivalent circuit of a static latch with ΔV_T; (b) Common-mode equivalent circuit; (c) Differential mode equivalent circuit.

2.2.3.1 Mismatched Trip Points

Let us denote the nominal trip point of each inverter as V_T. This is the point on its static I/O characteristic at which the input and output voltages are equal. In a well-designed inverter, $V_T = \frac{1}{2}V_{DD}$. Suppose that due to random spreads in threshold voltages, the trip points of the two inverters are unequal. Without loss of generality, we ascribe a deviation $+\Delta V_T$ to the trip point of one inverter and $-\Delta V_T$ to the other. The metastable point of the latch comprising these mismatched inverters lies at the coordinates in the phase plane $V_{OC} = \frac{1}{2}V_{DD}$, $|V_{OD}| = 2\Delta V_T$. That is, the metastable point is translated from its nominal position.

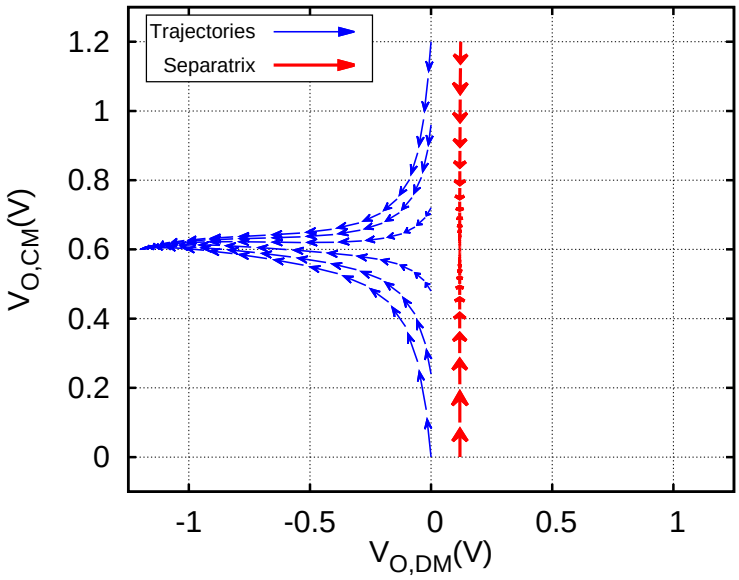


Figure 2.5: Phase plane of a static latch with V_T mismatch

Next we show how this mismatch will affect the vector field on the entire phase plane. First we identify the *separatrix* trajectory that leads into the metastable point. This we can do with the aid of equivalent circuits. To the first order the transconductance G_m of the two inverters remains matched for small ΔV_T. Now the linear equivalent circuit is as shown in Fig. 2.4. When the outputs are separated into modes, this circuit is equivalent to two circuits, one for the common mode and the other for the differential mode. The main point is that $2\Delta V_T$ appears *only* in the circuit for the differential mode an *independent* current source. By definition, the separatrix is that trajectory on the phase plane along which the unstable mode is *not stimulated*. When the unstable mode is stimulated and results in a growing exponent, in this circuit equivalent circuit the negative resistor $(-1/G_m)$ exchanges energy, or interacts, with a capacitor. But suppose the capacitors C_L were precharged (with the appropriate sign) to $2\Delta V_T$. Then the independent current source would find a return path through the two controlled sources, and no current flows through the capacitors. The unstable mode is not excited. Therefore, if the circuit is released from an initial condition with any common-mode voltage but with a differential voltage of $2\Delta V_T$, the common mode will decay into the metastable point and the differential voltage will remain constant for all time. In short, the separatrix that in the balanced circuit was vertical and coincident with

the axis $V_{OD} = 0$ is now, in the presence of mismatched trip points, *translated* by $2\Delta V_T$. An offset of this amount has appeared in the circuit Fig.2.5.

2.2.3.2 Mismatch in G_m

Random spreads in FET β will cause mismatch in the transconductance of the two inverters. Writing the transconductances as $G_m \pm \Delta\frac{1}{2}G_m$, we use equivalent circuits to examine the effects of this mismatch. We will assume that the trip points are matched, which means that the metastable point remains at the same location on the phase plane as for the balanced circuit.

In the equivalent circuit (Fig. 2.6(a)), this mismatch introduces an error current $\Delta\frac{1}{2}G_m V_{OC}(t)$ connected in a way that is itself clearly *not* in common-mode. The error current source belongs in the differential mode circuit. Removing this current source restores symmetry to the common-mode circuit, and it is readily seen that if the circuit is initialized with some common mode voltage $V_{OC}(0)$, this will decay into the metastable equilibrium value with a time constant set by the real pole $s_C = -G_m/C_L$.

The current owing to the mismatch ΔG_m appears in the differential mode equivalent circuit as an *independent* source because it is not affected by any of the variables in this circuit. It acts to *cross-couple* the modes [31]. This source's waveform must follow the decay of the common mode. We ignore the small perturbation of ΔG_m on $-G_m$. The differential mode's natural response is unstable, caused by a negative conductance charging the shunt capacitance $C_L + 2C_P$. However if at every instant the independent current source carries exactly the sum of the currents through the resistor and capacitor, then they will not interact and the unstable mode will not be excited. The states of the circuit will follow the separatrix into the metastable point. This requires that the differential voltage is initialized to a $V_{OD}(0)$ such that

$$\left[-G_m + s_C(C_L + 2C_P)\right] V_{OD}(0)\mathrm{e}^{+s_c t} = \Delta G_m V_{OC}(0)\mathrm{e}^{+s_c t}. \tag{2.4}$$

Assuming $2C_p \ll C_L$, this condition relates the initial conditions:

$$\frac{V_{OD}(0)}{V_{OC}(0)} \simeq \frac{\Delta G_m}{2G_m} \;\Rightarrow\; V_{OD}(0) \simeq \frac{\Delta G_m}{2G_m} V_{OC}(0) \tag{2.5}$$

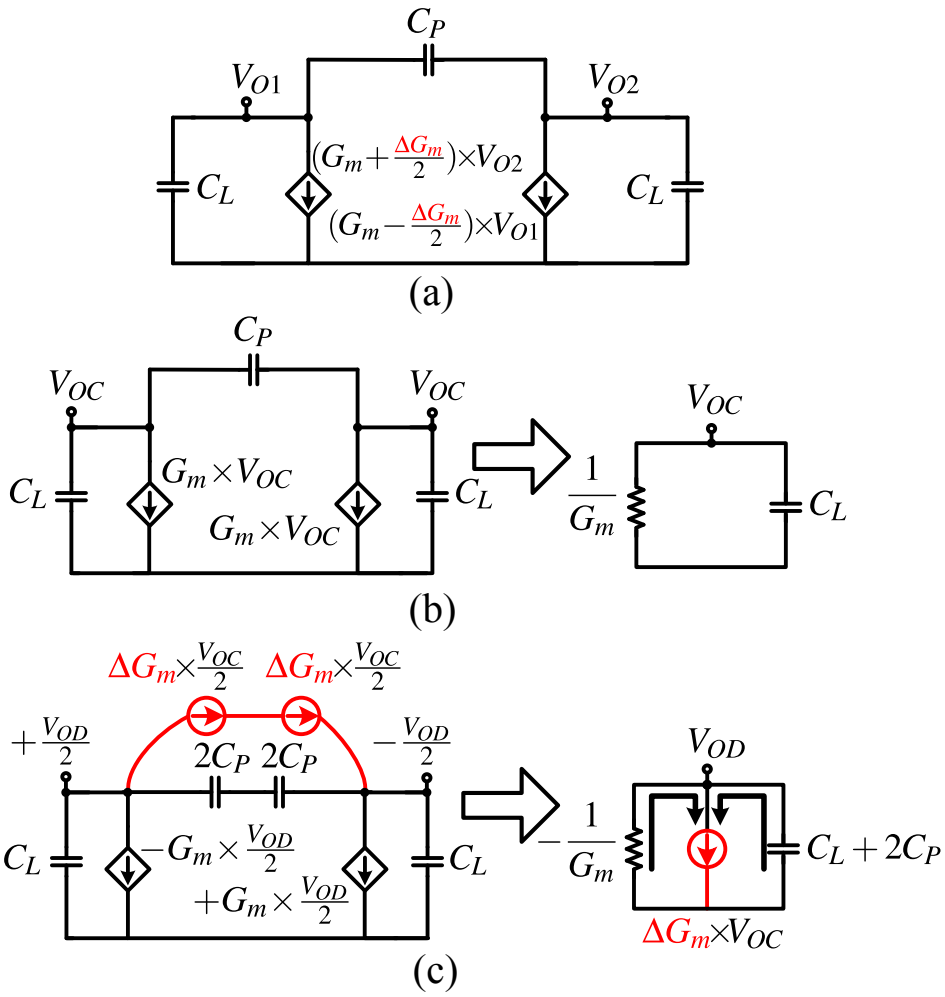


Figure 2.6: (a) Equivalent circuit of a static latch with ΔG_m; (b) Common-mode equivalent circuit; (c) Differential mode equivalent circuit with coupled source from common mode circuit.

If the output nodes are initially pulled up to the supply voltage, then $V_{OC}(0) = V_{DD}$. For the latch not to regenerate to '1' or '0', an initial differential voltage given by (2.5) must be applied at the same time. *This is the offset voltage caused by mismatch in* G_m. It is called a *dynamic* offset because it changes with the initial common mode V_OC0 forced at reset, which here depends on the supply voltage. On the phase plane, this means that the separatrix is *rotated* from a vertical line (Fig. 2.7).

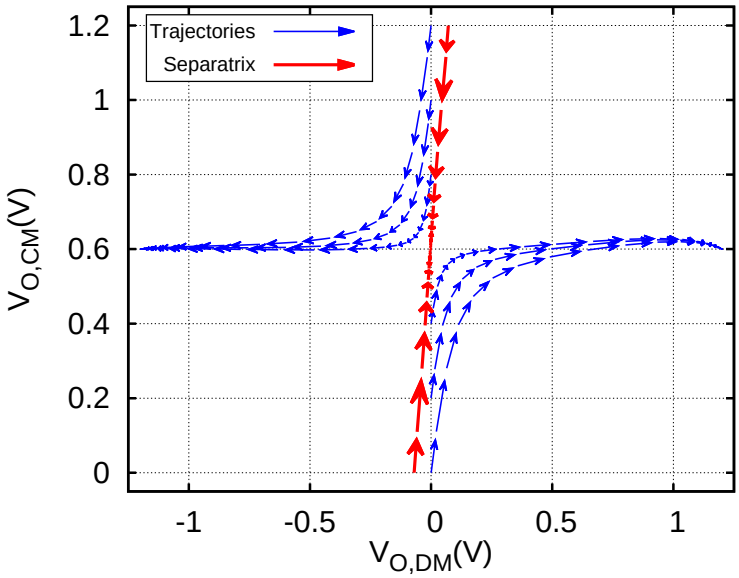


Figure 2.7: Phase plane of a static latch with G_m mismatch

2.2.3.3 Mismatch in Capacitors

Suppose all corresponding pairs of FETs are matched, but the load capacitors are mismatched. This is modelled by unequal capacitors $C_L \pm \frac{1}{2}\Delta C_L$ connected to the two outputs.

Again, this capacitor unbalance will introduce a current source that cross-couples the common mode waveform into the differential mode with a magnitude of $s\Delta C_L V_{OC}(s)$. Its effect, just like for G_m unbalance, is to rotate the separatrix, and to introduce an offset which depends on the initial value of the common-mode voltage as determined by the reset action:

$$\frac{V_{OD}(0)}{V_{OC}(0)} \simeq \frac{\Delta C_L}{2C_L + 2C_P} \;\Rightarrow\; V_{OD}(0) \simeq \frac{\Delta C_L}{2C_L + 2C_P} V_{OC}(0) \tag{2.6}$$

This is another dynamic offset. (2.6) is consistent with [15], but here this result is arrived at much more straightforwardly.

2.2.4 Offset Compensation

For small unbalances, the static offset and the two dynamic offsets may be treated independently, and the net offset due to all three is the algebraic sum of the independent offset. This raises the question of whether it is possible, using some method of adjustment, for this

sum to be forced to zero. The easiest offset to adjust would be that caused by ΔC_L, since it is routine practice in CMOS circuits to use arrays of small binary weighted capacitors that are digitally switched for calibration. So, as shown in Fig. 2.10, when trip point unbalance translates the metastable point and G_m mismatch rotates the separatrix, then there exists some ΔC_L of appropriate magnitude and sign that will also rotate the separatrix so that it passes through the coordinate $(0, V_{DD})$ on the phase plane. This means that if the latch nodes are reset to V_{DD}, then with zero differential input the latch will arrive at its metastable point. That is, by introduction of the correct capacitor unbalance with a digitally controlled array, the latch will sense and regenerate a differential input *free of all offset*.

This is very useful, but in the static latch we are analyzing Fig.2.1 it suffers from a practical flaw. As the supply voltage changes in a mixed-signal circuit during operation, the calibrated offset will depart from zero and no calibration loop will be able to track fast changes in V_{DD}. Another latched comparator circuit is needed whose offset is inherently resistant to changes in the supply voltage. It would be a bonus if its offset is mainly determined by one pair of FETs, to simplify calibration and lead to reliable offset-free operation. The so-called StrongArm latch, introduced in the next section, is such a circuit.

2.3 Dynamic Amplifiers

To analyze the class of comparators in widespread use today we must first understand the properties of a *dynamic amplifier*. All comparators use implicit dynamic preamplification before they regenerate, except for the memory sense circuit which for reasons of extreme compactness employs the simple cross-coupled CMOS latch.

A dynamic amplifier is a low-power circuit that amplifies a static input voltage by converting it into a current, then integrating that current on a capacitor over a well-defined time window [32, Sec. 5]. Since a comparator detects only the sign of a (small) input voltage, the preamplifier need not be linear, only that it should preserve the sign of the input—this means that its own offset and noise should be small—and that it should scale up the input voltage sufficiently to overcome noise and offsets contributed by the transistors that realize

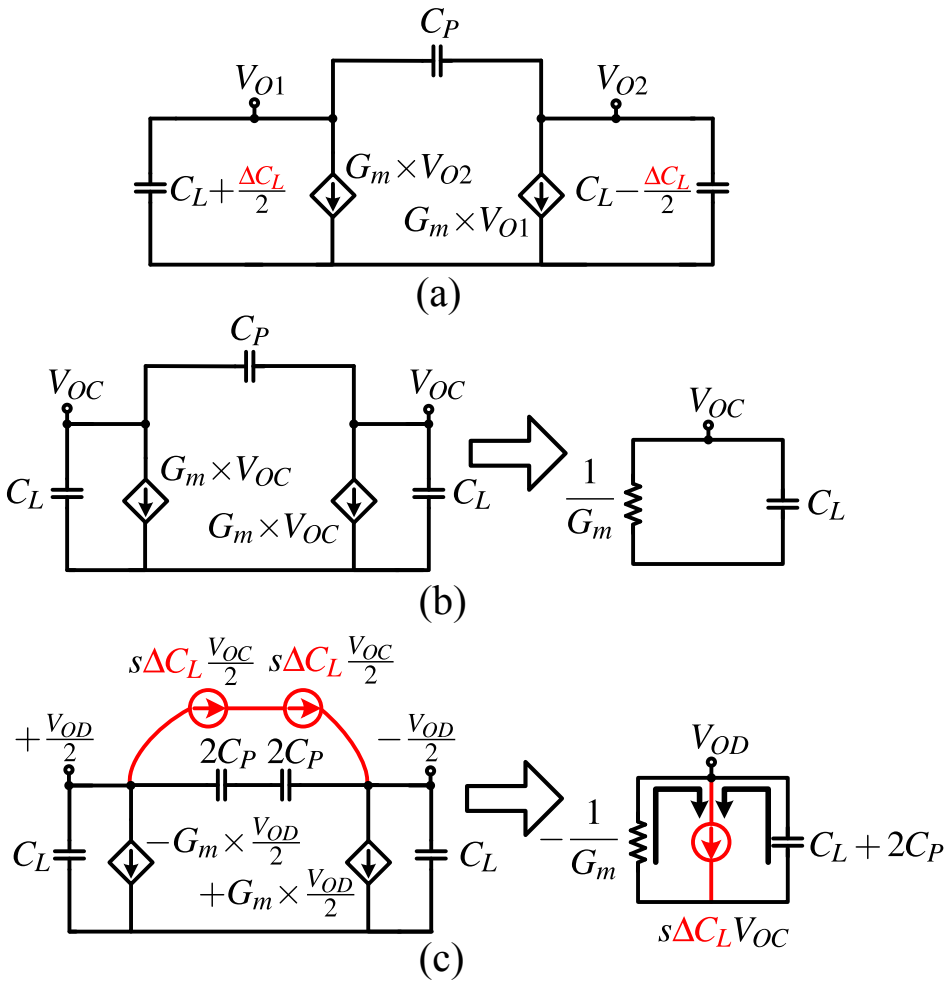


Figure 2.8: (a) Equivalent circuit of a static latch with ΔC_L; (b) Common-mode equivalent circuit; (c) Differential mode equivalent circuit with coupled source from common mode circuit.

the regenerative latch that follows. If offsets and noise in the regenerative latch are to be analyzed, [22] gives a framework to do so.

2.3.1 Gain

For our purposes, the simplest model of a balanced dynamic amplifier is a differential pair biased with current I_0, with equal grounded capacitors C attached to the drains. Two switches pre-charge the capacitors to the supply voltage V_{DD}. While a static differential voltage $v_{id} = V_{G1} - V_{G2}$ is applied, the switches are opened and amplification starts. This

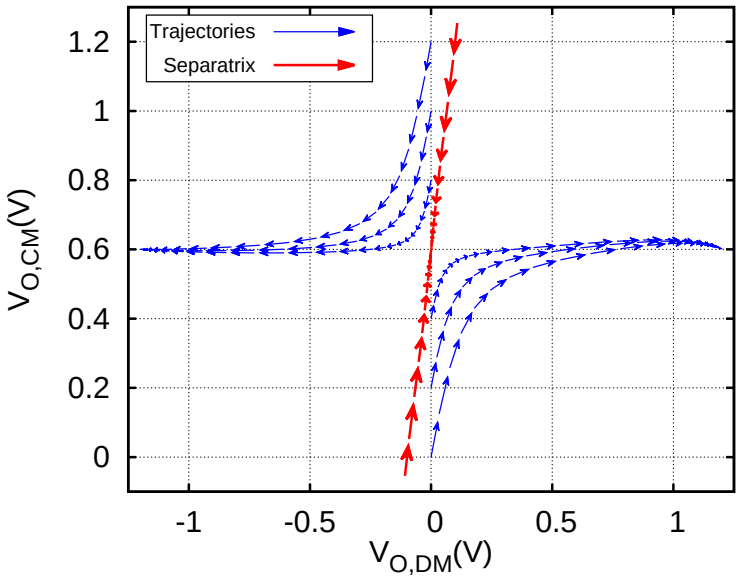


Figure 2.9: Phase plane of a static latch with C_L mismatch

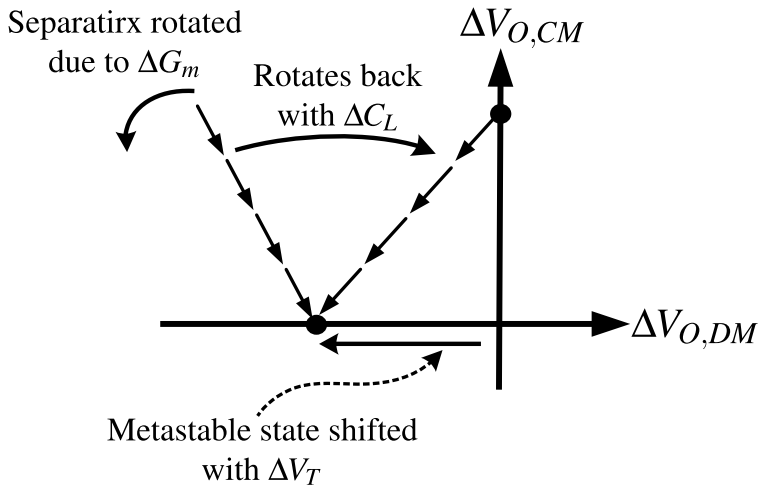


Figure 2.10: Phase plane of a static latch with offset calibration.

phase lasts for a time window t_W defined by the time required for the average (common-mode) voltage $\frac{1}{2}(V_{O1} + V_{O2})$ at the two drains to fall by some predetermined voltage V_t. Thus,

$$t_w = \frac{CV_t}{1/2 I_0}. \tag{2.7}$$

Suppose each FET is sized so that it needs a minimum voltage V_{DSAT} to operate in saturation. Then $g_m/I_0 = 1/V_{DSAT}$, where g_m is associated with a single FET. Over t_w, the

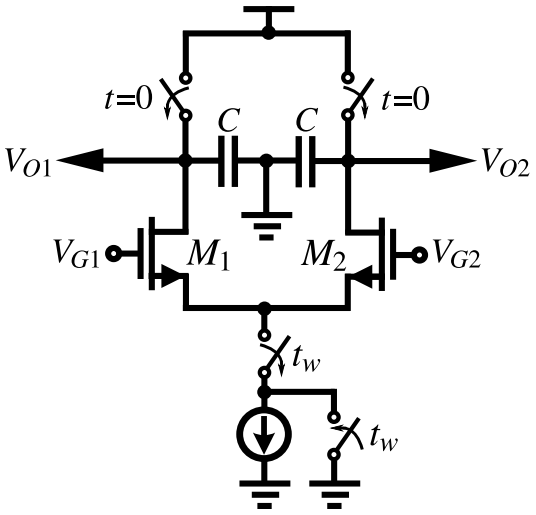


Figure 2.11: Conceptual dynamic amplifier.

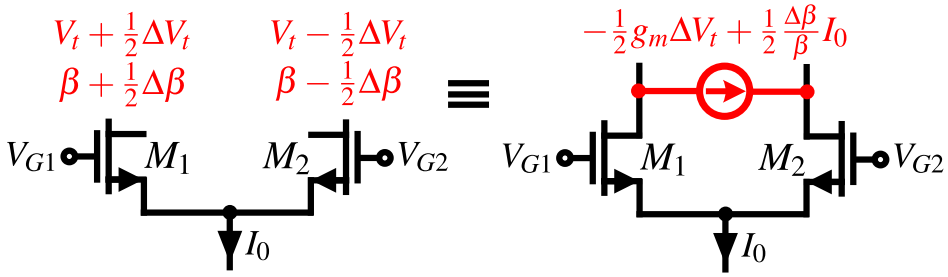


Figure 2.12: Modeling mismatch pair of FETs as balanced circuit with internal differential current.

differential current $\frac{1}{2}g_m v_{id}$ integrates on the differential capacitance $\frac{1}{2}C$ to create a differential output voltage [32, (40)]

$$v_{od} = \frac{g_m {}^1\!/_2 v_{id} t_w}{{}^1\!/_2 C} = \frac{2V_t}{V_{DSAT}} v_{id} \tag{2.8}$$

This is true when the FETs operate in saturation throughout. This simple model assumes that by some means, I_0 will cease to flow for $t > t_W$ when the amplification phase is complete. Then the gain of the dynamic amplifier is therefore the ratio of two voltages, one of which, V_{DSAT}, may be designed.

2.3.2 Circuit Unbalances

We will show that in a well-designed comparator, transistor mismatches in the input stage of the dynamic preamplifier determine the overall offset. To calculate these offsets, we must develop models of small mismatches between the symmetric pairs of circuit elements comprising the balanced topology of a dynamic amplifier. Each source of mismatch can be analyzed separately and, as long as they are small, their effects on the output added algebraically; in other words, superposition applies. We follow Middlebrook [31] in modelling mismatches with differential-mode currents that would flow in an otherwise perfectly balanced circuit free of mismatch, the currents being controlled by common-mode quantities. The threshold voltage and current scaling factor of each MOSFET are random variables, with mean value V_{t0} and β and standard deviations σ_{Vt} and σ_β. Let the threshold voltages of M1, M2 be V_{t1}, V_{t2}. Then $1/2(V_{t1} + V_{t2}) \simeq V_{t0}$, and

$$V_{t1} = V_{t0} + 1/2\Delta V_t;\ V_{t2} = V_{t0} - 1/2\Delta V_t;\ \Delta V_t = V_{t1} - V_{t2} \tag{2.9}$$

Given the MOSFET square-law characteristics,

$$I = \tfrac{\beta}{2}\left(V_G - V_{t0} - V_S\right)^2 = \tfrac{\beta}{2}V_{DSAT}^2 \tag{2.10}$$

$$\Rightarrow g_m = \beta\left(V_G - V_{t0} - V_S\right) \tag{2.11}$$

it follows that when equal V_G is being applied to the FET pair M1,M2, then after including mismatch in V_t,

$$I(M1) = 1/2(I_0 - g_m\Delta V_t) \tag{2.12}$$

$$I(M2) = 1/2(I_0 + g_m\Delta V_t) \tag{2.13}$$

These expressions are captured by an equivalent circuit (Fig. 2.12(b)) consisting of a *perfectly matched* differential pair with V_G applied to both gate terminals, each FET conducting $1/2I_0$ *and* a current source of $-1/2g_m\Delta V_t$ attached between the two drain terminals. In this way, mismatch is modelled by a perfectly balanced circuit with no differential stimulus applied to its input, that is, driven in common mode only, but with an independent differential current interpolated into the circuit that is proportional to mismatch ΔV_t. In other words,

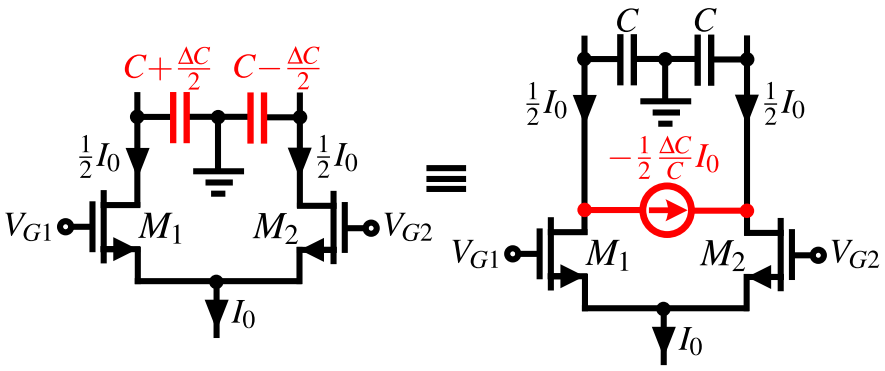


Figure 2.13: Modelling capacitor mismatch in a balanced circuit.

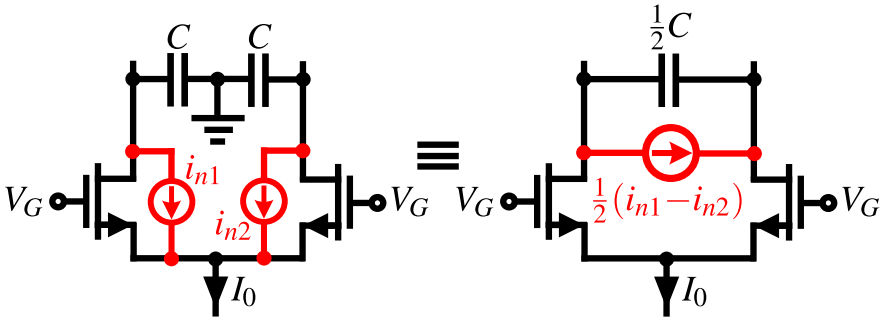


Figure 2.14: Noise in a balanced differential pair.

in a balanced circuit topology that is biased purely in common mode, unbalance due to component mismatch can be thought of as "cross-coupling" [31] the common-mode input into a differential-mode stimulus internal to the circuit. This coupling will become very clear when we consider capacitor mismatch later.

Along the same lines, the mismatch $\Delta\beta = \beta_1 - \beta_2$ induces another component in the differential current above, whose value is $+1/2(\Delta\beta/\beta)I_0$ (Fig. 2.12(b)).

Now suppose that the FETs are perfectly matched, but only the load capacitors are mismatched as $C \pm \frac{1}{2}\Delta C$. When the differential pair is biased with equal V_G, the two FETs will carry equal currents $1/2 I_0$; but these two currents flowing through mismatched load capacitors will create voltage ramps of unequal rates and thus a differential voltage that grows with time. This is the same as if the capacitors are perfectly matched, and are being charged by a differential current source $-1/2(\Delta C/C)I_0$ between the drains of the FETs. This shows most clearly how component mismatch is equivalent to the common-mode current I_0

cross-coupling into a proportional differential-mode current that appears in the appropriate location within a perfectly matched, balanced circuit.

The three contributions to the differential current, all assumed small, will superpose as

$$-\tfrac{1}{2}g_m\Delta V_t + \tfrac{1}{2}(\Delta\beta/\beta)I_0 - \tfrac{1}{2}(\Delta C/C)I_0 \tag{2.14}$$

(2.14) shows that offset is compensated by a *fractional unbalance* in capacitance. Does this mean that C can take on any value? As we will show, C sets the noise in the dynamic amplifier. Thus the circuit can be designed for specifications on offset and noise independently.

2.3.3 Noise in Dynamic Amplifier

Noise in the balanced dynamic amplifier originates in the two FETs. We will ignore flicker noise for now, which can be modelled as a small but slowly time-varying offset that cannot always be calibrated.

White noise occupying a very wide bandwidth may be modelled as an independent current source i_n between the source and drain of a noiseless FET, with spectral density $S_{in} = 4kT\gamma g_m$. With two FETs connected as a differential pair, noise appears as a differential current $i_{nd} = \frac{1}{2}(i_{n1} - i_{n2})$ inserted into the same circuit branch as the differential current that models mismatch. Noting that $i_{n1,2}$ are independent, the spectral density of this differential current is $S_{in,d} = 2kT\gamma g_m$.

The dynamic amplifier operates over a time window t_W, when its amplified output is passed to the next circuit stage. While the signal current is being integrated, so is wideband white noise. But the signal is static over the window t_W, whereas the noise current can fluctuate rapidly. This is the classic integrate-and-dump receiver. At the end of one window, the noise will have integrated on the capacitor to some random voltage. The integrated noise voltage sampled at the end of many such windows will follow a distribution whose variance, or mean square value, is given by [33, p. 331]

$$\langle v_0^2\rangle = S_{in,d}\frac{t_w}{C^2}\ (\mathrm{V}^2) \tag{2.15}$$

In answer to the question posed at the end of the last section, the nominal C determines mean-square noise voltage.

2.4 StrongArm Latch

This widely used latching comparator circuit in Fig. 2.15 was originally presented as part of a suite of low-power digital circuits [16]. It offers a convenient method to introduce a voltage to be regenerated into a pair of cross-coupled inverters—the main weakness of the CMOS static latch—while guaranteeing zero static power consumption when regeneration is complete. It gained widespread attention after it was used in the StrongARM microprocessor.

A survey of the literature suggests that in spite of widespread use, the detailed action of the StrongARM latch is, even now, poorly understood. Without a full understanding it cannot be used properly as a low offset comparator. Therefore, we will first explain how, in correct operation, the circuit traverses two phases over which the applied voltage is amplified before regeneration starts. When poorly designed, an internal regeneration can be triggered as early as in the second phase, but with the undesired consequence of a worse offset.

2.4.1 Overall Operation

The difference between two input voltages, each measured with respect to ground, is coupled into the latch through the NMOS pair M1-M2. The pair is activated by the tail FET, M_{CLK}, which is assumed to act like a switch. The input voltage common-mode (V_{IC}) must lie above the threshold voltage of M1,M2 and sets the initial bias current.

This bias current also flows through the cross-coupled inverters, M3-M5 and M4-M6, that are stacked in series. Differential current produced by M1-M2 "unbalances" the inverter, causing it to regenerate on this unbalance. Regeneration forces one FET in each inverter to turn OFF, thus choking off a current flow path through both M1 and M2. Thus M1, M2, and the tail current FET are all forced into deep triode with $V_{DS} = 0$ where they no longer conduct current. The comparator consumes no static power in its regenerated state.

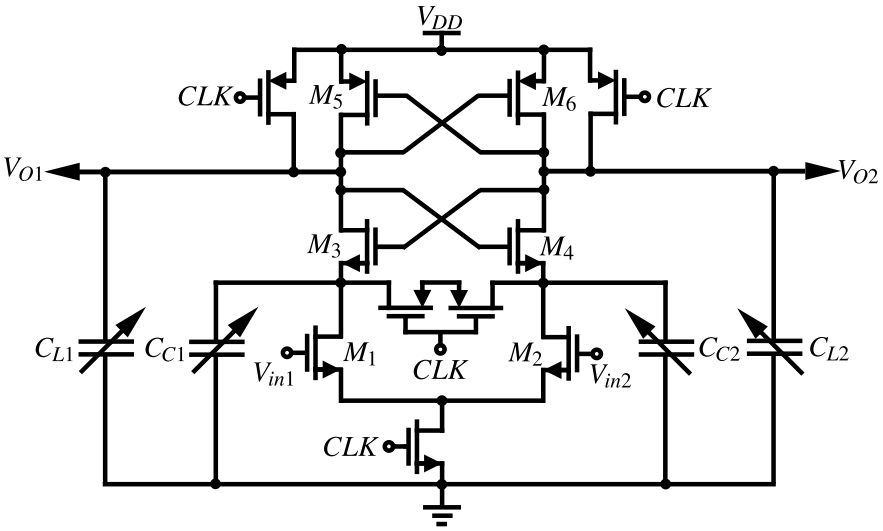


Figure 2.15: Schematic of the StrongARM latch.

This description is sufficient to see why the StrongARM latch is popular. A closer analysis is needed to understand aspects such as the circuit's inherent latency before it regenerates, and how unbalances in the circuit elements will cause static and dynamic offsets.

2.4.2 Operational Phases

The circuit's operation should be divided into three discrete phases Fig.2.16, with regeneration understandably taking place in the last phase. In the first two phases, sampling and propagation, the circuit amplifies the applied differential voltage on to internal nodes, as we now explain.

2.4.2.1 Reset State

The phases are most clearly identified when the comparator is released from a well-defined state. The circuit is defined by four state variables, the voltages on two grounded capacitors C_C and on two load capacitors C_L. But it is a time-varying circuit, and when in the second and third phases the capacitors exchange charge, the number of states collapses to two. Be that as it may, the circuit must be initialized with all four states at a predetermined and fixed value so as to erase memory of the previous regeneration. A convenient initialization

is to precharge all four capacitor voltages through FET switches to the supply voltage V_{DD}. This initializes the source and drain terminals of M3-M6 all to the same potential.

2.4.2.2 Sampling Phase

The input voltages V_{in1} and V_{in2} are applied when the tail current transistor M_{CLK} is initialized. The average input voltage sets the bias current (I_0) through M1 and M2. This common-mode current can initially only discharge C_{C1} and C_{C2}. M3 and M4 will remain OFF, until the capacitors have discharged by an amount equal to the threshold voltage V_{tN}. This period of time defines the sampling phase,

$$T_s = \frac{C_C \cdot V_{tN}}{I_0} \tag{2.16}$$

Over this period the difference in the input voltages v_{ID}, which creates a differential current $i_{ID} = g_{m1,2} v_{ID}$, integrates a differential voltage across the capacitors C_C. This is best seen in a differential half circuit as shown in Fig.2.16(b)

$$V_{CD,s} = \frac{i_{ID} \cdot T_s}{C_C} \tag{2.17}$$

This differential voltage serves as the initial condition for the next phase, propagation.

2.4.2.3 Propagation Phase

In propagation phase, M3 and M4 will turn ON, and by the end of this phase the common mode (bias) current flowing through them will have discharged both output voltages V_{O1} and V_{O2} by $|V_{tP}|$ to turn on the cross-coupled PMOS pair M5, M6. During the propagation phase M1,M2 and M3,M4 are ON. The voltages on capacitors C_C and C_L will ramp down together, separated by the constant difference of $V_{GS3} (= V_{GS4})$. Thus,

$$T_p = \frac{(C_L + C_C)|V_{tP}|}{I_0} \tag{2.18}$$

Although the gates of M3,M4 are cross-coupled to the drains, in common mode the pair of gates follows the same voltage waveform as the pair of drains, as if each FET was diode connected. The cross-coupling only becomes evident in differential mode. A cross-coupled

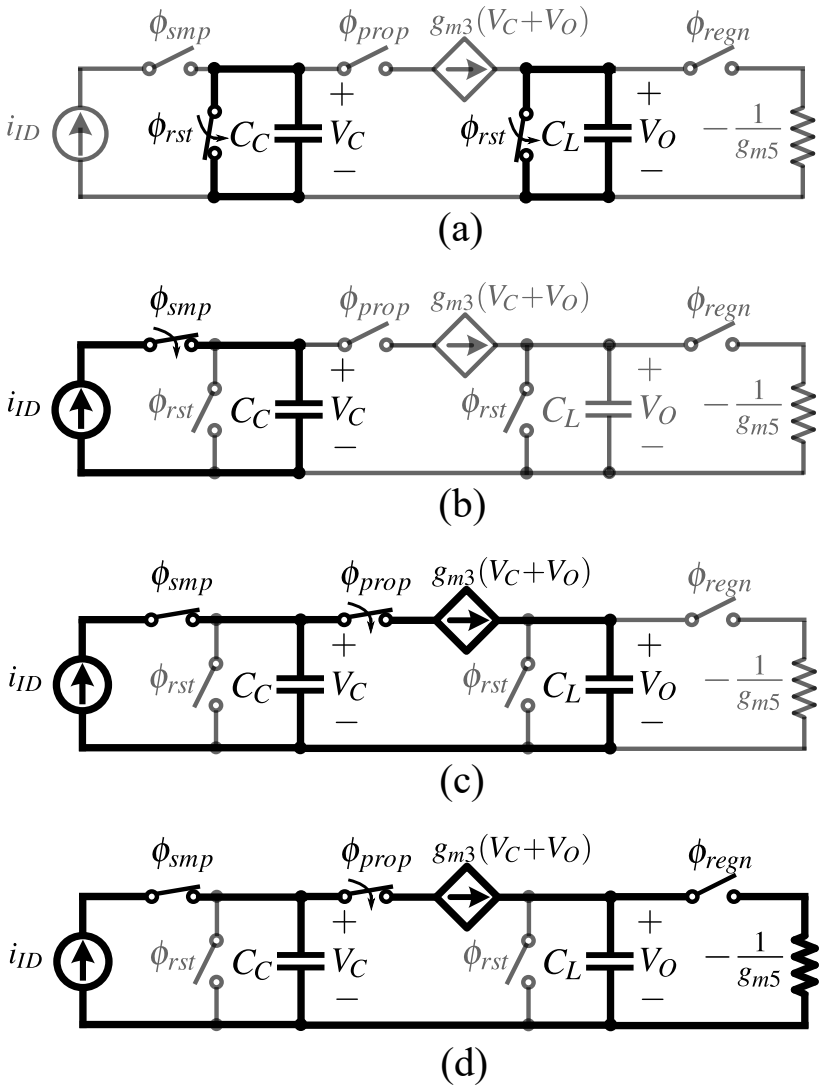


Figure 2.16: Differential half circuit of the StrongARM latch during (a) reset phase; (b) sampling phase; (c) propagation phase; (d) regeneration phase.

pair of transistors will regenerate if the loop gain is greater than one. This is not always so when, as we have assumed, M1,M2 remain in saturation through the propagation phase, and $C_C < C_L$. This is contrary to [34], which claims a guaranteed right half plane pole as long as M3,M4 conduct. This is readily proved by replacing M3,M4 with a transconductance g_{m3} in the linearized equivalent circuit of Fig. 2.16(b), and representing M1,M2 with a constant current source. The cross-coupling shows in the control variable of g_{m3}, which instead of being the familiar difference of two voltages, each measured with respect to ground, is now their sum. The voltages V_C and V_O are the differential voltages across C_C and C_L, respectively.

The two capacitors in this equivalent circuit are in series and the circuit has a single pole at

$$s_p = -\frac{g_{m3,4}(C_L - C_C)}{C_L C_C} \tag{2.19}$$

where $g_{m3,4}$ is defined by the common mode current through M3,M4

$$g_{m3,4} = \frac{2I_0}{V_{P3,4}} \cdot \frac{C_L + C_C}{C_L} \tag{2.20}$$

Depending on the relationship between C_C and C_L, the pole in the differential equivalent circuit can be in either left or right half plane. In most practical cases, C_L comes from the loading capacitors and C_C comes from drain capacitors of MOSFETs. Usually $C_L > 2C_C$ applies. This condition keeps the pole in the left half plane and the differential circuit acts as a stable amplifier. The differential current from M1,M2 is amplified by the cross-coupled NMOS M3,M4 into C_L. Sometimes C_C is programmed as a calibration capacitor arrays to correct offset, when C_C can approach C_L or become even the dominant capacitor. When $C_C > C_L$, the pole in the differential circuit crosses into right half plane, indicating regeneration. However, regeneration only occurs over the time window before cross-coupled PMOS pair begins to conduct. It still behaves as an amplifier, with a gain determined by regeneration within this window [35]. Irrespective of ratio between C_C and C_L, we need to characterize the output voltage on C_L and the preamplification from the input to the output.

The differential voltage on C_C during sampling phase propagates to C_L through charge sharing, resulting in a differential voltage on C_L of

$$V_{OD1} = \frac{i_{ID} \cdot T_s}{C_C} \frac{C_C}{C_L - C_C} \left(1 - \mathrm{e}^{+s_p T_p}\right) \tag{2.21}$$

Meanwhile, provided M1,M2 remain in saturation, the differential current between M1,M2 integrates on C_C and C_L, with the differential current onto C_L boosted by the cross-coupled NMOS pair M3,M4. The differential current charging C_L through propagation phase is

$$i_{OD}(t) = i_{ID} \cdot \frac{C_L}{C_L - C_C} \left(1 - \mathrm{e}^{+s_p t}\right) \tag{2.22}$$

The differential voltage due to current integration (T_P) over the duration of the propagation

phase is then calculated by

$$\begin{aligned} V_{OD2} &= \int_0^{T_p} \frac{i_{OD}(t)}{C_L}\,\mathrm{d}t \\ &= \int_0^{T_p} \frac{i_{ID}}{C_L - C_C}\left(1 - \mathrm{e}^{+s_p t}\right)\,\mathrm{d}t \end{aligned} \tag{2.23}$$

Both charge sharing and current integration contribute to the differential voltage on C_L, thus the net differential voltage is the superposition of the two

$$\begin{aligned} V_{OD} &= V_{OD1} + V_{OD2} \\ &= i_{ID}\left(\underbrace{\frac{T_s + T_p - \tau_p}{C_L - C_C}}_{\text{linear integration}} + \underbrace{\frac{\tau_p - T_s}{C_L - C_C}\cdot \mathrm{e}^{+s_p T_p}}_{\text{exponential}}\right) \end{aligned} \tag{2.24}$$

where $\tau_p = -1/s_p$ represents the time constant of the differential circuit during propagation phase. When $\tau_p < 0$ with large C_C, the differential circuit becomes regenerative. Because the exponential component varies significantly depending on the ratio between C_C and C_L, it is necessary to quantify the ratio of T_p/τ_p,

$$s_p T_p = -\frac{T_p}{\tau_p} = -\frac{2V_t}{V_{P3,4}}\cdot\frac{C_L - C_C}{C_C} \tag{2.25}$$

When $C_C < C_L/2$, $\exp(+s_p T_p) \ll 1$ and the exponential term can be ignored. The circuit is simply a linear integrator. When $C_C \approx C_L$ and $\exp(+s_p T_p) \simeq 1$, Taylor expansion $e^x \approx 1 + x$ is used for simplification . When $C_C > 2C_L$, the regenerative exponential term $\exp(+s_p T_p)$ is much larger than 1 and dominates over the linear integration. With these simplifications, simple forms of (2.24) are obtained for physically more intuitive estimations,

$$V_{OD} = i_{ID} \times \begin{cases} \dfrac{T_s + T_p}{C_L - C_C} & \text{if } C_C < C_L/2 \\ \dfrac{T_s + \frac{1}{2}T_p}{C_L - C_C}\cdot\dfrac{T_p}{\tau_p} & \text{if } C_C \approx C_L \\ \dfrac{\tau_p - T_s}{C_L - C_C}\cdot \mathrm{e}^{+s_p T_p} & \text{if } C_C > 2C_L \end{cases} \tag{2.26}$$

When $C_C - C_L < 0$, τ_p is also negative, still rendering positive output voltage V_{OD}. The transimpedance gain A_G through sampling and propagation phase is defined as $A_G = V_{OD}/i_{ID}$

2.4.2.4 Regeneration Phase

At the end of the propagation phase, the cross-coupled PMOS pair, M5,M6, turns ON and, since its source terminals are shorted, it will regenerate the differential voltage present at the output nodes Fig.2.16(c). This voltage is, to a very good estimate, given by (2.26). The pole for PMOS pair is must lie in the right-half s-plane, and is given by

$$s_{reg} = +\frac{g_{m5}}{C_L} \tag{2.27}$$

The regenerated output voltages will grow to control all the FETs in the entire latch. At one of the two stable equilibria, one output voltage will reach V_{DD} and the other ground. No static current will flow on either side of the circuit. If the PMOS pair were to dwell at the metastable equilibrium and the supply voltage was large enough, then the output voltages would be equal, $V_{O1} = V_{O2} = V_{DD} - V_{GS}(M5, 6)$.

2.4.2.5 StrongARM Latch in Summary

This detailed analysis shows that the StrongARM latch, with all internal nodes initialized to V_{DD}, undergoes two phases of common-mode discharge before it regenerates. These phases amount to a delay of $(T_s + T_p)$ before the circuit regenerates with the time constant given by the inverse of the pole frequency (2.27). Others have noticed that the StrongARM latch regenerates after some latency, or waiting period [1] and consider this a weakness. But our analysis shows that over this period of latency a useful internal amplification takes place. For the applied input voltage v_{ID}, the gain to instant of regeneration is

$$A_V = \frac{V_{OD}}{v_{ID}} = \begin{cases} \dfrac{C_L + 2C_C}{C_L - C_C} \cdot \dfrac{2V_t}{V_{P1,2}} & \text{if } C_C < C_L/2 \\ \dfrac{C_L + 3C_C}{C_C} \cdot \dfrac{V_t^2}{V_{P1,2}V_{P3,4}} & \text{if } C_C \approx C_L \\ \dfrac{2V_t + V_{P3,4}}{V_{P1,2}} \cdot \exp\left(\dfrac{2V_t}{V_{P3,4}}\right) & \text{if } C_C > 2C_L \end{cases} \tag{2.28}$$

where $V_{P1,2}$ is the pinchoff, or overdrive, gate voltage on M1,M2. By substituting typical values into the two terms above, this gain is about $5 \sim 15$. [36] correctly identifies the phases

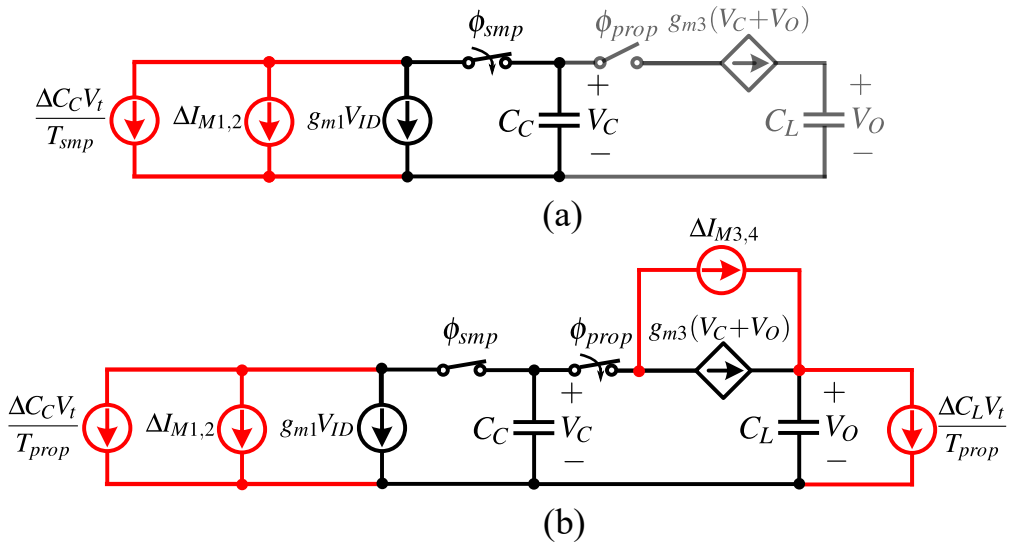


Figure 2.17: Equivalent differential circuit model of strongARM latch with mismatch sources during (a) sampling phase and (b) propagation phase.

of the strongARM latch, but fails to recognize the effect of cross-coupled NMOS pair and underestimates the preamplfication gain.

2.4.3 FET Mismatch

Estimating impacts of different FET mismatches requires understanding transfer functions from each mismatch source to the comparator output. Shown in Fig.2.17 are the equivalent differential mode circuits with potential mismatches during sampling and propagation phases. Since mismatch sources are assumed to be small perturbations on the operation of a balanced circuit, superposition applies and each FET mismatch can be studied independently.

Both V_t and β mismatch in the input pair M1,M2 appear at the comparator input as an effective differential offset voltage.

$$V_{OS}|_{M1,2} = \Delta V_{t1,2} + \frac{\Delta\beta_{1,2}}{\beta_{1,2}} \cdot \frac{V_{P1,2}}{2} \tag{2.29}$$

The input common-mode 'kickback' caused by the pulse of gate charge when CLK turns on M_{CLK} will be converted to a differential offset voltage by β mismatch [37]. This offset needs to be reduced or calibrated in all applications. Another 'kickback' comes when the latching of M5,M6 causes one of M1,M2 to turn off and the other to be forced into triode.

Gate charge is expelled by FET turning on, and a small amount of charge flows into the other FET. This is different from the common-mode kickback in both origin and impact. The differential 'kickback' only happens after the strongARM latch finishes regeneration and does not interfere the decision itself. In flash ADCs which use resistor reference ladders, this kickback must be reduced [38] as the ladders' voltages will ring due to this coupling [39, CH. (8)]. But in charge redistribution SAR ADCs with only one comparator making a decision every cycle, the differential-mode kickback is not important [40].

Mismatch in the PMOS cross-coupled pair M5,6 contributes negligibly to the input referred offset because when M5,6 conduct, the internal preamplification has typically amplified the input signal enough to overcome mismatch in M5,6.

Threshold mismatch between M3,4 introduces a constant current $\Delta I_{M3,4}$ parallel to the controlled source during propagation phase as in Fig.2.17(b). This current does not change the total charge on C_C and C_L, and does not contribute significantly to the output voltage. As currents through M3,4 are set by M1,2, β mismatch between M3,4 does not result in differential current between M3,4. Thus, the contribution from M3,4 to the input referred offset is negligible.

We have proved that the internal gain amplifies the imbalances in the input differential pair most, usually to the point that they will dominate all other FET imbalances in the circuit. This identifies the principal source of offset, and design can focus on its mitigation. In a well designed strongARM comparator, input referred offset is dominated by M1,2 alone; sizing M3-6 changes it little.

2.4.4 Capacitor Mismatch

C_C and C_L, both grounded capacitors, are vulnerable to mismatch. To model their contribution to dynamic offset in the simplest way, we follow the method shown in the static latch analysis where capacitor mismatch induces a coupling from common-mode to differential-mode Fig.2.17. Since the StrongARM latch operates by discharging both C_C from their reset condition by V_{tN} over the sampling mode, and then, by discharging C_L and C_C by $|V_{tP}|$

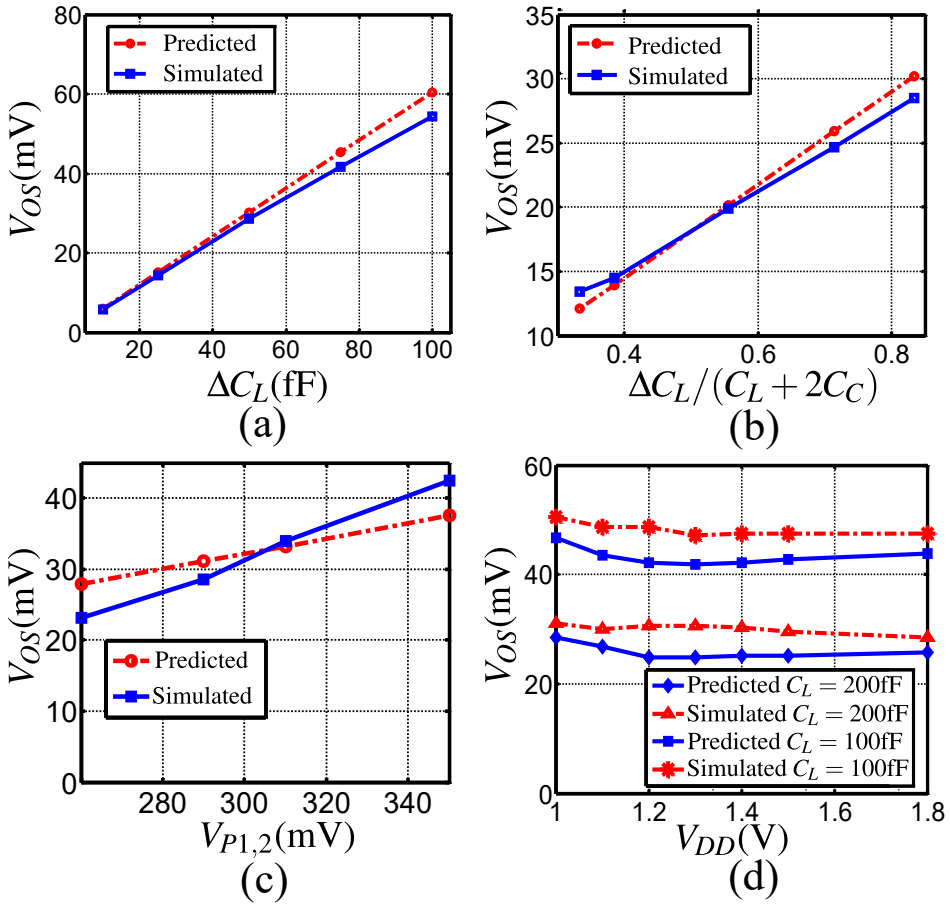


Figure 2.18: Comparison between theoretical calculation and simulation: (a) C_L = 200fF, C_C = 20fF, $V_{P1,2}$ = 250mV, vary ΔC_L from 10fF to 100fF; (b) C_L = 200fF, ΔC_L = 50fF, $V_{P1,2}$ = 250mV, vary C_C from 20fF to 200fF; (c) C_L = 200fF, C_C = 20fF, ΔC_L = 50fF, vary $V_{P1,2}$ from 260mV to 350mV; (d) ΔC_L = 50fF, C_C = 20fF, C_L = 100fF and 200fF respectively, vary V_{DD} from 1V to 1.8V.

over the propagation phase, we expect small mismatch between corresponding pairs of these capacitors to induce significant (differential) mismatch over the large excursions of common mode.

Through sampling and propagation phase, the common-mode to differential-mode coupling current due to C_C mismatch integrates charge onto C_C and C_L. The total charge induced by C_C mismatch is $\Delta C_C(V_{tN} + |V_{tP}|)$. To prevent the circuit from regeneration when the PMOS pair become active, an differential offset voltage must be applied to the input so that differential voltage at the end of propagation phase is zero. This requires the

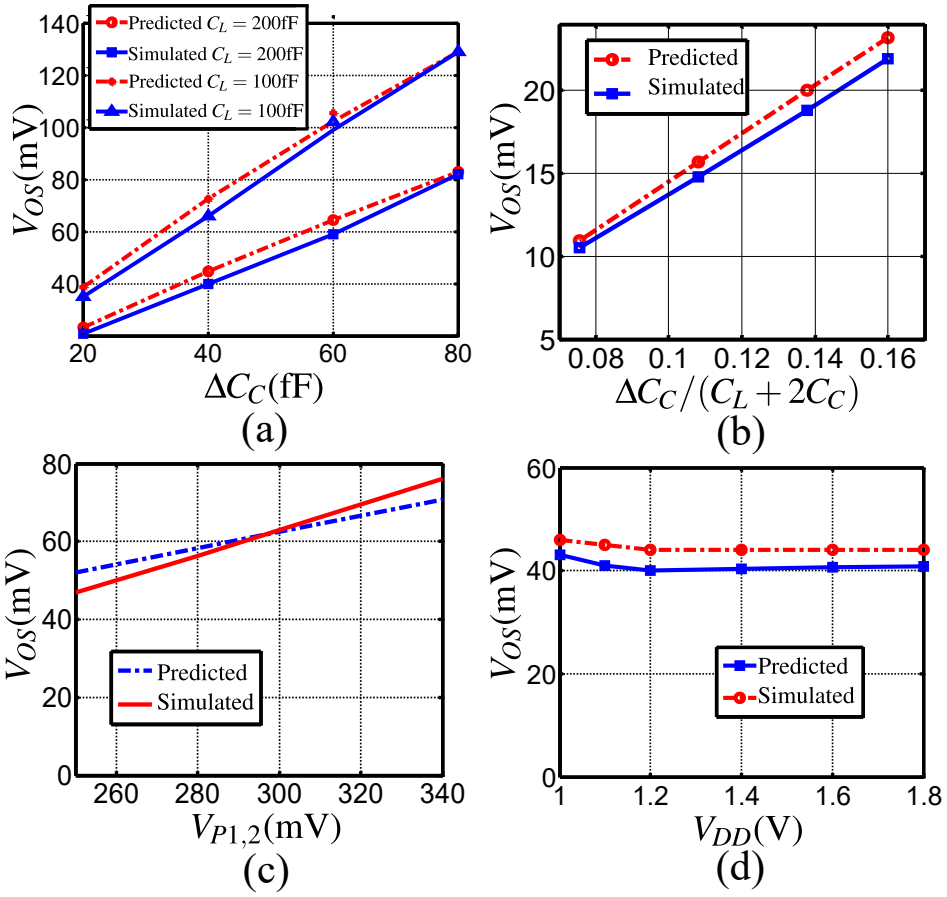


Figure 2.19: Comparison between theoretical calculation and simulation: (a) $C_L = 200$fF, $C_C = 40$fF, $V_{P1,2} = 250$mV, vary ΔC_C from 20fF to 80fF; (b) $C_L = 200$fF, $\Delta C_C = 50$fF, $V_{P1,2} = 250$mV, vary C_C from 20fF to 160fF; (c) $C_L = 200$fF, $C_C = 20$fF, $\Delta C_C = 20$fF, vary $V_{P1,2}$ from 250mV to 340mV; (d) $\Delta C_C = 20$fF, $C_C = 20$fF, $C_L = 200$fF, vary V_{DD} from 1V to 1.8V.

offset voltage contributing the same amount of charge onto C_C and C_L, thus from (2.16) and (2.18)

$$\begin{aligned} g_{m1,2}V_{OS}(T_s + T_p) &= \Delta C_C(V_{tN} + |V_{tP}|) \\ \Rightarrow V_{OS} &= \frac{2\Delta C_C}{C_L + 2C_C} \cdot \frac{V_{P1,2}}{2} \end{aligned} \tag{2.30}$$

Contrary to [21], (2.31) shows that the offset introduced by ΔC_L is independent of V_{DD}. While the common mode to differential-mode coupling induced by C_C mismatch lasts through both sampling and propagation phase, C_L mismatch induces a similar coupling only during

propagation phase, at the output port parallel to C_L. Similarly, estimating offset requires applying an input differential voltage to counter balance the common-mode to differential-mode coupling charge injected by ΔC_L into C_C and C_L.

$$g_{m1,2}V_{OS}(T_s + T_p) = \Delta C_L|V_{tP}|$$
$$\Rightarrow V_{OS} = \frac{\Delta C_L}{C_L + 2C_C} \cdot \frac{V_{P1,2}}{2} \tag{2.31}$$

C_C imbalance induces twice the offset of C_L imbalance because C_C undergoes a common-mode voltage excursion of $V_{tN} + |V_{tP}|$, almost twice the excursion of the voltage on C_L. Fig.2.19 and Fig.2.18 demonstrates the simulated dynamic offsets against estimated offsets in a strongARM latch in 90nm CMOS process. This circuit rejects typical bounce in supply voltage, so its regeneration time, static offset, and dynamic offset all remain independent of V_{DD}.

2.4.5 Offset Compensation

With this background, we can explain comprehensively, how the offset calibration strategies first proposed in [41] for latched comparators will work. We have established that with zero differential input, mismatch in the threshold voltage and β of the pair M1,M2 will usually create the offset voltage at the output nodes according to (2.28), dominating contributions from mismatch in all other FETs. Now if the capacitances C_C and C_L are fine tuned in closed loop under digital control to create an almost equal offset but opposite in sign according to (2.30)(2.31), the algebraic sum of the two offsets will cancel. So by introducing a deliberate and measured mismatch in the capacitors, the comparator may be made to appear offset-free even with random mismatch in the FETs .

The process is readily automated by applying a common-mode voltage to the two inputs of the comparator (zero differential input), which, due to inherent mismatches will cause the comparator to repeatedly produce a '0' or a '1'. The binary-weighted arrays of small capacitors attached to C_C and C_L are searched until the output of the latch changes state. That setting can be held in a register dedicated to each latch.

The expression (2.30) also shows a benefit of resetting the initial voltages across C_C to

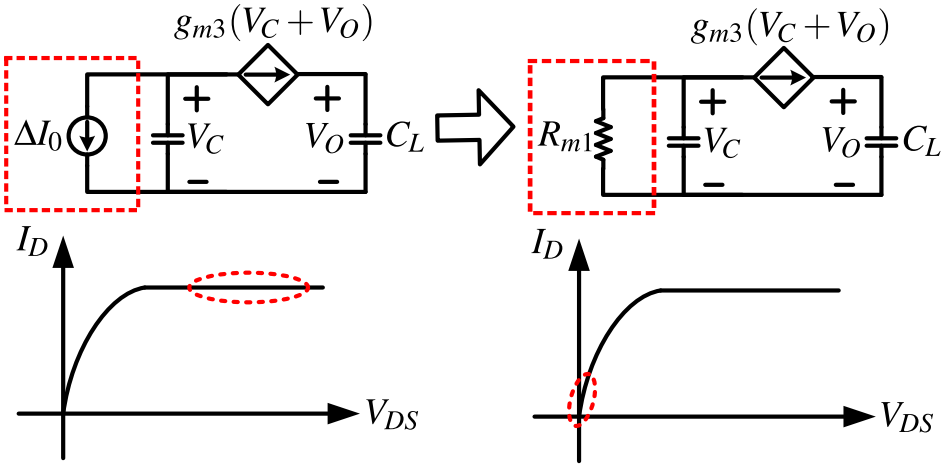
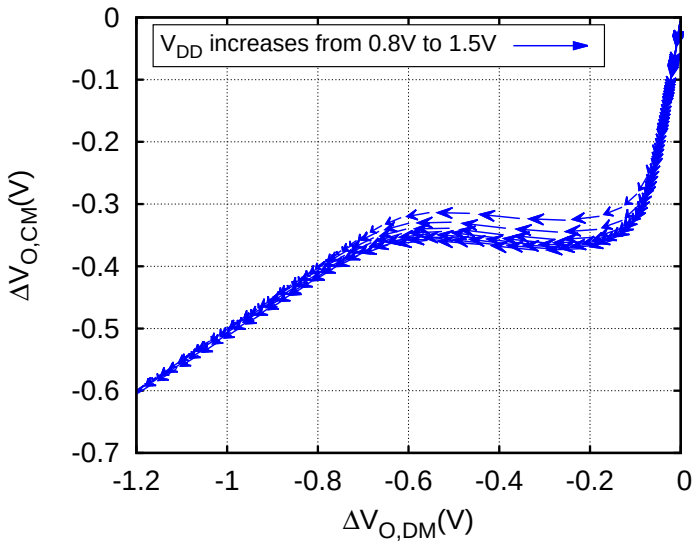


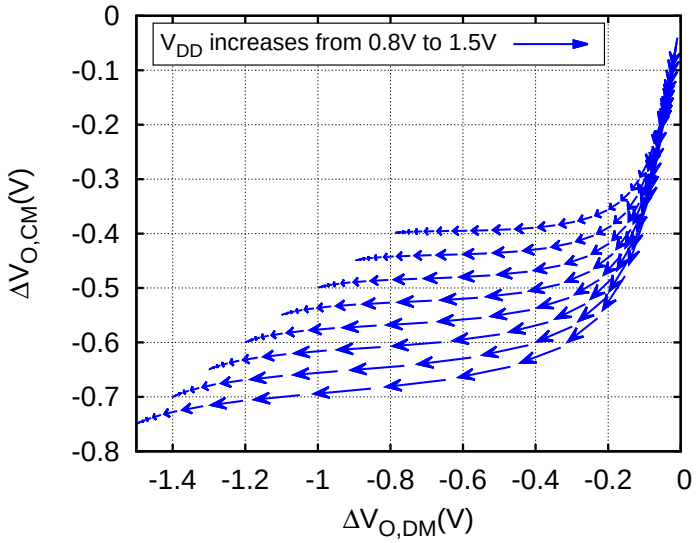
Figure 2.20: Equivalent differential circuit during propagation phase when $M1$ & 2 transit from saturation to triode region.

V_{DD}. When the internal nodes at the drains of M1,M2 are not reset, they are pulled up to $V_{DD}-V_{tN}$ when the output nodes are reset. Indeed, resetting the C_C nodes to V_{DD} prolongs the latency period by T_s (see (2.16)), which seems undesirable. But (2.30) reveals that as a result of this delay ΔC_C is twice as effective in compensating offset. Thus, at the expense of a larger latency, offset is compensated with a lighter capacitor loading.

But doesn't the greater latency annul the benefit of a lighter capacitor loading? To answer this question, we examine the critical comparator in an A/D converter. A comparator is in critical condition when it is resolving an input so close to a threshold that regeneration may not complete in the allotted clock phase; that is, the converter would make a metastability error. This is limited by the regeneration time constant, seldom latency. In the StrongARM latch it is the PMOS pair M5,M6 that regenerates, and in the initial part of the regeneration transient it is loaded only by C_L. When C_C shoulders the larger burden of compensating latch offset, the latch regeneration time constant is essentially unchanged. When the offsets are large, a better balance between latency and regeneration may require that the calibration arrays are more evenly distributed between the C_C and C_L nodes.



(a)



(b)

Figure 2.21: Phase plane plot of a strongARM latch with C_L mismatch, V_{DD} being swept from 0.8V to 1.5V, (a) M1, M2 remain in saturation; (b) M1, M2 enter triode during propagation phase.

2.4.6 Range of Input Common Mode

Poorly chosen input common mode can degrade offset and noise considerably for the reasons that we will now explain. This was observed experimentally in an early use of the latch as

an SRAM sense amplifier [1] and subsequently recognized by [2][36].

Over the sampling and propagation phases, the input common-mode voltage (V_{IC}) sets the bias current through M1-M4, but also determines whether or not these FETs operate in saturation. This is best understood in a perfectly balanced, offset-free latch with zero differential input, which upon release from reset should travel into its metastable state. It is sufficient to examine the state of the circuit when the regenerative pair M5,M6 starts to conduct, i.e. when $V_{O1} = V_{O2}$ falls from V_{DD} to $V_{DD} - |V_{tP}|$. Since this excursion is purely in common mode, M3,M4 will behave as if they are diode connected and therefore they will operate in their saturation region. But for the input pair M1,M2 to operate in saturation, V_{IC} must not exceed an upper limit

$$V_{tN} < V_{IC} < V_{DD} - V_{GS}(\text{M3,M4}) \tag{2.32}$$

The StrongArm latch dissipates zero static power because the analog input is coupled into the cross-connected inverters via a differential pair in series. But at low supply voltages, this stack of three FETs in series is easily driven into a poor operating region.

When the internal nodes are precharged to V_{DD}, the input pair will remain in saturation over the entire sampling phase for any V_{IC}, even if it is as large as V_{DD}. This ensures full dynamic amplification of the input voltage v_{id} on to the capacitors C_C over the interval T_s. Problems arise in the propagation phase, when M1,2 can be pushed into the triode region. There are two consequences to this: 1. M1,2 can discharge the amplified voltage on C_C: indeed, we will assume that they do so rapidly, and 2. The deeper in triode region that M1,2 operate, the lower their transconductance during the propagation phase, worsening the comparator's input-referred offset and noise. Possible erasure of the amplified voltage afflicts all comparators with dynamic pre-amplification and, as we will show, new comparator topologies announced in the literature offer, in effect, remedies to this problem.

Analysis of a StrongArm latch is messy and cumbersome when the input FETs operate in the triode region. But it is possible to simplify analysis for the special case when the gates of the input NFETs are biased at $V_{IC} = V_{DD}$. This is the highest voltage that can be applied to the gates, and it is certain to force the input FETs into triode during propagation.

This bias condition has been investigated experimentally for an SRAM sense amplifier, and measured offsets for different V_{IC} are reported [1].

The analysis to follow uses approximations which we now make explicit. The latch circuit is enabled with a small input v_{id}. The sampling phase proceeds normally as the precharged drain voltages V_{c1}, V_{c2} discharge by V_{tN}, the input FETs operate in saturation, and the dynamically amplified voltage appears on C_C. As the propagation phase commences, the input FETs are pushed into triode. Now we assume that 1. C_C is discharged immediately by the on conductance of the FETs in triode (this is an oversimplification); 2. an equal current flows through M1 and M3 in series (and through M2 and M4); and 3. the current through the FETs at the *start* of propagation can be used to predict how the circuit behaves *throughout* propagation.

Initially the gates and drains of M3,M4 are precharged to V_{DD}, that is, the gates of M1∼M4 are all at the same potential. It can be shown by using the square-law model of the FET I-V characteristic that if two NFETs with the same β are in series, their gates are tied together at voltage V_G, and the upper NFET (M3) operates in saturation, then the lower NFET (M1) *must* be in triode and its drain voltage is $(1-1/\sqrt{2})(V_G - V_{t0}) \simeq 0.3(V_G - V_{t0})$.

To measure the deterioration due to improper bias, we use as a reference the properly operating StrongArm latch. Because of dynamic amplification in a circuit with a well-chosen input bias V_{IC}, only the mismatches in V_t and β of the input pair contribute to offset ($v_{os,0}$),

$$\boxed{v_{os,0} = \Delta V_{t,12} + \frac{\Delta\beta_{12}}{\beta}\frac{I_D}{g_m} = \Delta V_{t,12} + \frac{\Delta\beta_{12}}{\beta}\frac{V_{IC} - V_{t0}}{2}.} \tag{2.33}$$

If, on the other hand, $V_{IC} = V_{DD}$, then using the expression given above for the drain voltage in triode region, the differential current due to mismatch in β of M1,M2 changes to

$$\begin{aligned} I_{D1} - I_{D2} &= {}^1\!/_2\Delta\beta_{12}\,(0.3(V_{DD} - V_{t0})(2V_{IC} - 0.3V_{DD} - 1.7V_{t0})) \\ &= {}^1\!/_2\Delta\beta_{12}(0.3 \times 1.7)(V_{DD} - V_{t0})^2. \end{aligned} \tag{2.34}$$

We have assumed that the preamplified differential voltage on C_C has been erased, so threshold mismatch in M3,M4 ($\Delta V_{t,34}$) will now unbalance the drain voltages of M1,M2 in triode,

causing

$$I_{D1} - I_{D2} = \beta(V_G - 0.3V_{DD} - 0.7V_{t0})\Delta V_{t,34}$$
$$= 0.7\beta(V_{DD} - V_{t0})\Delta V_{t,34}. \quad (2.35)$$

To refer all these unbalances to the latch input, we assume that the FET pairs are matched and a differential offset voltage v_{os} is applied to the input on a bias of $V_{IC} = V_{DD}$. Then from the same expressions

$$I_{D1} - I_{D2} = 0.3\beta(V_{DD} - V_{t0})v_{os}. \quad (2.36)$$

By combining in magnitude (2.34) and (2.35), we use (2.36) to find the equivalent input offset

$$v_{os} = \Delta V_{t,12} + \frac{1.7}{2}(V_{DD} - V_{t0})\frac{\Delta\beta_{12}}{\beta_{12}} + \frac{7}{3}\Delta V_{t,34}. \quad (2.37)$$

Comparing with (2.33), the contribution of β mismatch in the input pair scales up by a factor of 1.7×. So if V_{DD} =1.5 V and V_{t0} =0.3 V, this offset is almost 3× higher. The offset due to V_t mismatch in M3,M4, which was previously negligible, is now almost 2× larger than the offset due to V_t mismatch in M1,M2.

This analysis is meant to give insight, and is simplistic in places. For example, the amplified signal charge on C_C is not erased instantly, but a portion of it is transferred to C_L. As a result, (2.37) overestimates offset. Appendix 2.13 gives a more accurate analysis, which is used to calculate the entries in Table 2.2 that appears later in the paper.

2.5 Thermal Noise

Comparators are mainly limited by random or systematic offsets. Long-standing methods to circumvent offsets such as overranging and digital error correction in multi-step A/D converters are now up against limits posed by low supply voltages. In ADCs with resolution above 10bits, thermal noise from comparator usually dominates over the kT/C noise in track/hold circuit. Moreover, the input referred root-mean-squared (RMS) noise of a strongARM latch

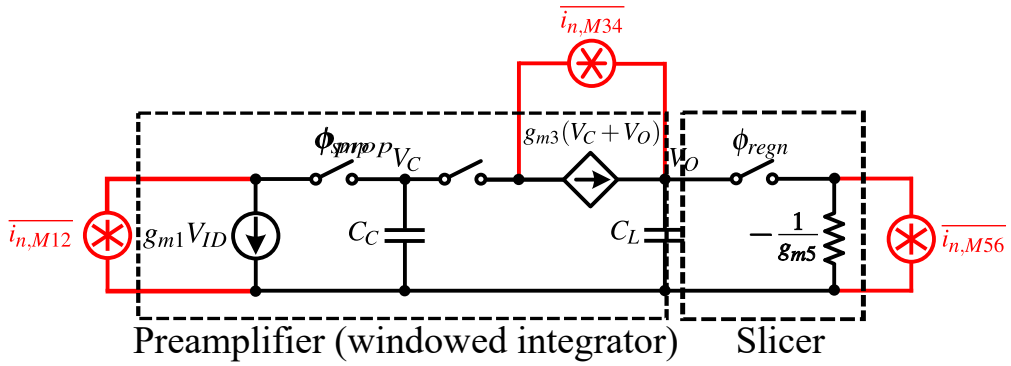


Figure 2.22: Equivalent differential circuit model for noise calculation.

may vary 2.5 times across PVT corners [42]. Before resorting to sophisticated techniques such as noise shaping [43, 44] or statistical estimation [42] to reduce this noise, we must understand its origin and answer the fundamental question: how serious is it in limiting the comparator?

There is some published work to model noise in the StrongARM latch [19, 20]. [19] identifies the operational phases of this circuit correctly, but it uses a method of noise analysis that is too indirect to yield key insights: for instance, the finding that the input common mode must be chosen correctly (see Section 2.4.6 above) in order to minimize noise.

On the other hand, our analysis of offset in the StrongARM latch extends straightforwardly to thermal noise Fig.2.22. But first we must describe how noise will randomly trigger regeneration around the metastable point. In a perfectly balanced comparator with zero input applied, the noise current in the input pair M1,M2 will integrate on C_C during the sampling phase, and continue to integrate and amplify on C_L during the propagation phase. The random voltage integrated on C_L will trigger regeneration to '0' or '1'. Other noise sources from M3-M6 do not contribute significantly as described in Section 2.4.2. Here we limit our discussion to strongARM latch for its wide usage, [45] discusses noise in a static latch alone.

The output noise can be obtained either using integration on frequency domain or time domain. Suppose a linear transfer function has the transfer function $H(j\omega)$ and a corresponding impulse response $h(t)$, a white noise S_{in} at its input port results in an output noise

power of

$$\langle v_{n,out}^2 \rangle = S_{in} \int_0^{+\infty} |H(j\omega)|^2 \mathrm{d}\omega = S_{in} \int_0^{+\infty} h^2(t) \mathrm{d}t \tag{2.38}$$

This conclusion is proved by Parseval's theorem. In this dissertation, we choose to calculate noise using time domain integrations as it provides more concise and compact solutions. It is not necessarily so in many other cases, especially when the shape of PSDs is of concern [46].

The noise current from M1,M2 integrates on C_L during sampling phase, resulting a noise voltage on C_C which then propagates to C_L during propagation. The integrated noise voltage on C_C is the outcome of a noise current integrating a single capacitor C_C

$$\langle v_{n,c}^2 \rangle = S_{in} \int_0^{T_s} \left(\frac{1}{C_C} \right)^2 \mathrm{d}t = S_{in} \cdot \frac{T_s}{C_C^2} \tag{2.39}$$

This can be also obtained by integrating $|H(j\omega)|^2$ on frequency domain. For a windowed integrator

$$H(j\omega) = \frac{1}{j\omega C_C}(1 - \mathrm{e}^{-j\omega T_s}) \rightarrow |H(j\omega)|^2 = \frac{T_s}{C_C^2} \left(\frac{\sin \frac{\omega T_s}{2}}{\frac{\omega T_s}{2}} \right)^2 \tag{2.40}$$

The same output noise power is obtained using frequency domain integration

$$\langle v_{n,c}^2 \rangle = S_{in} \int_0^{+\infty} \frac{T_s}{C_C^2} \left(\frac{\sin \frac{\omega T_s}{2}}{\frac{\omega T_s}{2}} \right)^2 \mathrm{d}t = S_{in} \cdot \frac{T_s}{C_C^2} \tag{2.41}$$

$S_{in} = 4\mathrm{kT}\gamma g_{m1,2}$ represents the noise current PSD in M1,M2, where k is the Boltzmann constant, T is the temperature and γ is a process technology dependent constant, usually between $2/3 \sim 1$. At the end of propagation phase, the noise power on C_L due to charge sharing from $\langle v_{n,c}^2 \rangle$ is found by

$$\langle v_{n,out1}^2 \rangle = S_{in} \cdot \frac{T_s}{2C_C^2} \left[\frac{C_C}{C_L - C_C} \left(1 - \mathrm{e}^{+s_p T_p} \right) \right]^2 \tag{2.42}$$

As described in Sec.2.4.2.3, the noise current from M1,M2 also integrates on C_L during propagation phase. The transfer function, without considering the windowing operation, from input current to the output voltage is given by

$$H_0(s) = \frac{V_{OD}(s)}{I_{ID}(s)} = \frac{g_{m3,4}}{C_C C_L} \cdot \frac{1}{s - s_p} \tag{2.43}$$

Its corresponding impulse response is

$$h_0(t) = \frac{1}{C_L - C_C}\left(1 - \mathrm{e}^{+s_p t}\right) \tag{2.44}$$

With propagation phase truncating the effective integration time, the impulse response is finally

$$h(t) = \frac{1}{C_L - C_C}\left(1 - \mathrm{e}^{+s_p t}\right)[\mathbf{u}(t) - \mathbf{u}(t - T_p)] \tag{2.45}$$

Applying inverse Laplace transform to (2.45), the transfer function including the windowing operation is then found by

$$\begin{aligned} H(s) = \frac{g_{m3,4}}{C_C C_L} \cdot \frac{1}{s - s_p} + \frac{1}{C_L - C_C} \cdot \frac{1}{s} \cdot \mathrm{e}^{-sT_p} \\ - \frac{1}{C_L - C_C} \cdot \frac{1}{s - s_p} \cdot \mathrm{e}^{+sT_p} \end{aligned} \tag{2.46}$$

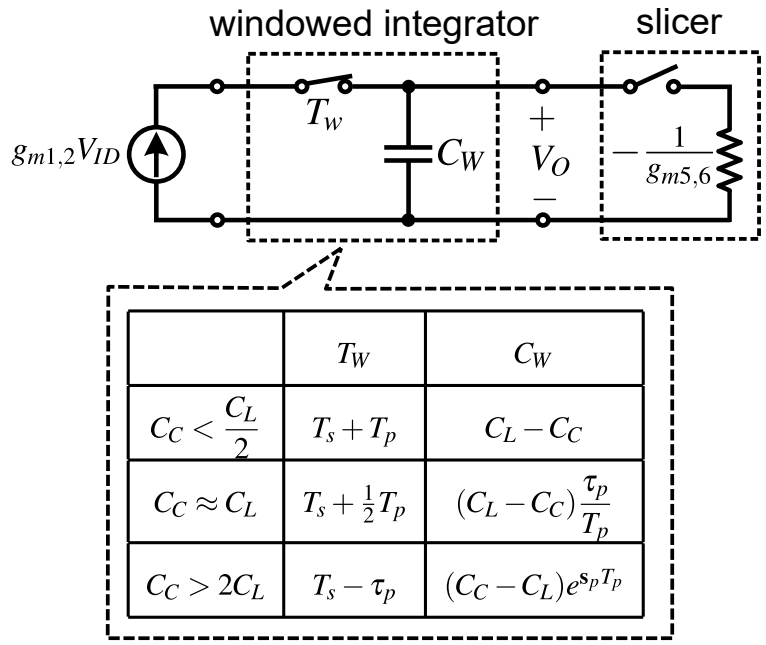
$|H(j\omega)|^2$ is further obtained by replacing s with $j\omega$ in (2.43). Comparing (2.46) and (2.45), $h(t)$ is a natural choice for calculation because the integration window is readily embedded. The contribution to the output voltage noise power is then found by

$$\langle v_{n,out2}^2 \rangle = S_{in} \int_0^{T_p} \left[\frac{1}{C_L - C_C}\left(1 - \mathrm{e}^{+s_p t}\right)\right]^2 \mathrm{d}t \tag{2.47}$$

The input referred current noise power is calculated by dividing the total output noise power $\langle v_{n,out}^2 \rangle = \langle v_{n,out1}^2 \rangle + \langle v_{n,out2}^2 \rangle$ with A_G^2 as calculated in (2.26),

$$\langle i_{n,in}^2 \rangle = \frac{\langle v_{n,out}^2 \rangle}{A_G^2} = S_{in} \times \begin{cases} \dfrac{1}{T_s + T_p} & \text{if } C_C < C_L/2 \\ \dfrac{T_s + \frac{1}{3}T_p}{(T_s + \frac{1}{2}T_p)^2} & \text{if } C_C \approx C_L \\ \dfrac{1}{T_s - \tau_p} & \text{if } C_C > 2C_L \end{cases} \tag{2.48}$$

The equivalent noise bandwidth (NBW) is defined as the bandwidth of a brickwall filter that filters a wideband white noise to produce the same integrated input/output noise. Thus,



	T_W	C_W
$C_C < \frac{C_L}{2}$	$T_s + T_p$	$C_L - C_C$
$C_C \approx C_L$	$T_s + \frac{1}{2}T_p$	$(C_L - C_C)\frac{\tau_p}{T_p}$
$C_C > 2C_L$	$T_s - \tau_p$	$(C_C - C_L)e^{\mathbf{s}_p T_p}$

Figure 2.23: The most simplified equivalent circuit model for a strongARM latch including both noise, preamplification gain and speed.

NBW for a strongARM latch is found by inspecting (2.48),

$$\text{NBW} = \begin{cases} \dfrac{1}{T_s + T_p} & \text{if } C_C < C_L/2 \\ \dfrac{1}{T_s + \frac{1}{2}T_p} & \text{if } C_C \approx C_L \\ \dfrac{1}{T_s - \tau_p} & \text{if } C_C > 2C_L \end{cases} \tag{2.49}$$

It indicates that when C_C is smaller than $C_L/2$, the equivalent NBW is entirely determined by the duration of sampling and propagation phase, which are then dictated by the discharging speed in the common-mode circuit and independent of the differential-mode circuit pole s_p itself. On the other hand, when C_C becomes substantially larger than C_L, which might happen in some cases where C_C is configured as calibration capacitor arrays, the equivalent NBW will be determined by both the sampling phase duration and the time constant in the differential-mode circuit.

Comparing the voltage gain and the NBW in a strongARM latch with a standard capacitor alone integrator, the most simplified equivalent circuit model for a strongARM latch is

then constructed as shown in Fig.2.23, consisting an capacitor-only integrator followed by an ideal slicer for regeneration. This simple circuit model captures both the voltage preamplification and the equivalent NBW of a strongARM latch. Since sampling phase provides efficient amplification that helps reduce both offset and noise Fig.2.23, the benefit of resetting internal nodes prevails the cost of the extra latency.

As the input referred noise voltage is of final interest for design, it is obtained with $\langle v_{n,in}^2 \rangle = \langle i_{n,in}^2 \rangle / g_{m1,2}^2$

$$\langle v_{n,in}^2 \rangle = 4\mathrm{kT}\gamma \cdot \frac{V_{P1,2}}{V_t} \times \begin{cases} \frac{1}{2} \cdot \frac{1}{C_L + 2C_C} & \text{if } C_C < C_L/2 \\ \frac{1}{C_L + 3C_C} & \text{if } C_C \approx C_L \\ \frac{1}{C_C(2 + V_{P3,4}/V_t)} & \text{if } C_C > 2C_L \end{cases} \tag{2.50}$$

Unlike dynamic offset that depends on the ratio of capacitor imbalance, the input referred noise can be reduced only by increasing the total capacitor size or reducing the bias voltage $V_{P1,2}$. This is a direct trade-off between noise and speed as C_C increases the latency before final regeneration and C_L impacts the regeneration time constant. As (2.50) reveals, increasing C_C is more effective in reducing noise than C_L. However, it is most efficient to keep C_C around $C_L/2$ as its efficiency in reducing noise decreases as C_C grows. Though a low common mode also reduces noise by lowering $V_{P1,2}$, it is limited by PVT variations. Unlike capacitors that have relatively stable values, even implemented as MOS capacitors [47, Eq. (15)], threshold variations across PVT corners varies $V_{P1,2}$ significantly. This results in large variations in noise as reported in [42], which is undesirable and poses additional design complexity.

(2.50) offers straightforward estimation for noise at the design stage so that designers can decide if it is indeed necessary to apply more sophisticated techniques to reduce noise. Furthermore, it helps designers make the practical design choice whether to apply an auxiliary differential pair for offset calibration or to utilize calibration capacitor arrays for the same purpose. In low-to-moderate resolution ADCs, when strongARM comparators satisfy noise requirements with intrinsic capacitance from transistors, auxiliary differential pairs correct

offsets without penalty in speed [2]. Yet, it should be noticed that the auxiliary differential pair itself introduces extra noise to the input. While in high resolution ADCs where noise budget is stringent, using capacitor arrays is preferred over auxiliary differential pairs because the calibration capacitors are able to correct offsets and reduce noise simultaneously. The calibration capacitors do not pose extra penalty as they are absorbed into capacitors indispensable to lower the thermal noise.

2.6 Speed

The net delay of a StrongArm latch consists of two parts: 1. the *latency* of the sampling and propagation phases, which is determined by the common mode bias current, and 2. the delay in *regeneration*, which is mainly set by the acceptable metastabilty error rate. We explain what the latter means. In any regenerative waveform, the final value at the end of a clock phase is proportional to the initial voltage. Since the input to a comparator is a random voltage uniformly distributed over a well-defined interval, there is a sub-interval around zero where the drive is too small for the comparator output to regenerate to a logic threshold within the clock window. The bit-error rate is given by the ratio of this sub-interval to the full interval [39, Sec. 8.1.3]. This dictates the minimum clock period as a multiple (n) of the regeneration time constant, plus latency. The latency is independent of the input differential signal, whereas the regeneration delay is affected by both the applied input voltage and the output voltage after amplification.

$$\boxed{t_{delay} = (T_s + T_p) + n\tau_{reg}} \tag{2.51}$$

While both C_L and C_C will determine the latency for common-mode discharge, the regeneration time constant $\tau_{reg} \approx C_L/G_m$, where G_m is the net transconductance of the cross-coupled inverters. In the initial part of the regeneration phase, cross-coupled pair M5,M6 drives C_L only while M3,M4 in saturation isolate it from C_C. Later, when M1 and M2 are pushed into triode, M3,M4 will add to G_m their own transconductance degenerated by M1,M2. C_C will be shorted by M1,M2. Thus, G_m grows over the regeneration phase. Expressions (2.30), (2.31), (2.50) and (2.51) together offer a quantitative guide to the optimum design of a

StrongArm latch for offset, noise and speed.

2.6.1 Kickback

With zero differential input, the gates of M1,M2 absorb a pulse of current to support inversion layers when M_{CLK} turns them on. In the critical comparator with near-zero differential input, this is essentially a common-mode current pulse extracted from the circuit connected to the comparator input; but β mismatch [37] in M1,M2 will convert it into a small differential current. Since the driving circuit is not a perfect voltage source, this converts into a differential offset voltage. The impedances driving the two inputs are often unequal, which convert the common-mode current into a differential offset voltage. In flash ADCs where one input may be connected to a resistor reference ladder, a distributed RC circuit, this kickback must be lowered [38] since the current impulse will induce ringing on tap voltages [39, Ch. 8] and cause reference errors, and if settling is slow, interference with the next decision. Another 'kickback' appears during regeneration, when the latching of the cross-coupled inverters M3∼M6 causes one of M1,M2 to turn off and the other to be forced into triode. Gate charge is expelled by the FET that turns off, and a small amount of additional charge flows into the other FET. This is different from the common-mode kickback in both origin and impact. The differential 'kickback' only happens after the StrongArm latch completes regeneration and therefore does not interfere with the decision itself. Charge-redistribution SAR ADCs, where one comparator makes every decision, are a special case. [40, Appendix D-D] shows that differential-mode kickback causes no harm as long as the comparator is reset before every decision.

2.7 Double Tail Comparator

The "double tail" comparator preserves the benefits of internal dynamic amplification in the StrongArm circuit at low supply voltages, or when the input common-mode bias is unsuitable. Originally developed to buffer the input circuit from charge kickback during regeneration [3], its circuit structure (Fig. 2.24(a)) also forces the static power to zero. A

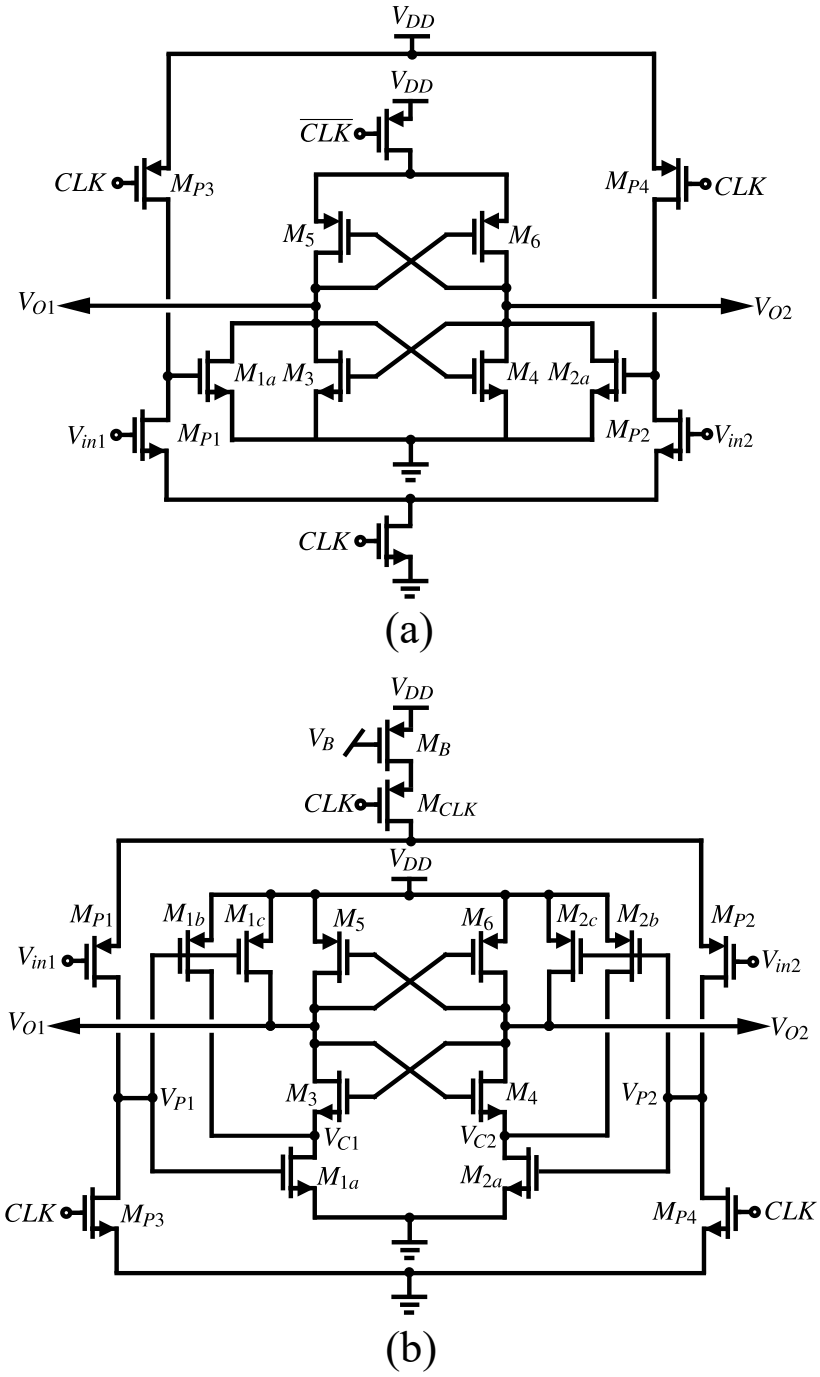


Figure 2.24: (a) The original double-tail latch proposed from [3]; (b) The improved double-tail latch proposed from [4].

dynamic preamplifier precedes a latch. When CLK goes high, the amplified differential input appears superposed on the output common-mode (bias) voltage that in this circuit ramps down from an initial voltage of V_{DD} to 0. This changing common-mode voltage will ultimate-

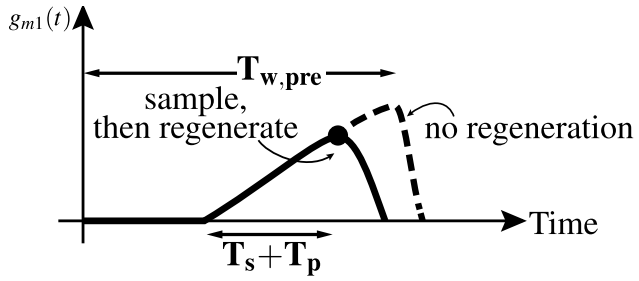


Figure 2.25: $g_{m1,2}(t)$ as a time dependent parameter in a double-tail comparator.

ly force the preamplifier differential pair MP1,MP2 into triode, which will then discharge the amplified differential voltage stored on the output capacitors, Thus the preamplifier will shut off its own bias current. In other words, signal amplification is available over the time window between when the clock turns on the amplifier, and before it shuts itself off, and erases the amplified voltage stored on the capacitors C_{pre}.

The FETs M1a,M2a that couple the dynamic preamplifier to a static latch are key to correct operation. Initially these FETs are biased into deep triode by the large voltage V_{P1}, V_{P2} applied to their gates. Their β is designed much larger than of M3,M4 to lower the loop gain of the cross-connected pair M5,M6, suppressing regeneration. While V_{P1}, V_{P2} ramp down, MP1,MP2 continue to amplify the differential input; at some point M1a,M2a release their clamp on M3,M4, enabling the latch M3$_\sim$M6 to regenerate. Ideally the sign of the amplified differential voltage coupled into the latch through M1a,M2a, will determine the regenerated binary output. Only the preamplifier is clocked in this circuit; it actuates the latch circuit.

[25] points out with simulations that this arrangement performs poorly. We will explain why. Consider the dynamics of the coupling FETs M1a,M2a. They turn on in deep triode with a large $V_G = V_{DD}$ and $V_D \approx 0$. As V_G ramps down, their V_D rises, until V_G crosses below V_{t0} when they turn off. This describes a time trajectory on the FET I_D-V_{DS} plane which starts in deep triode when the FETs' g_m is low; then the FET enters saturation, when V_D rises to $^1/_2V_{DD}$ but at that point $V_G \to V_{t0}$, so again their g_m will be low. This means that over this trajectory the transconductance is weak most of the time, so M1a,M2a are ineffective in coupling the preamplified voltage into the latch. In fact they worsen input-

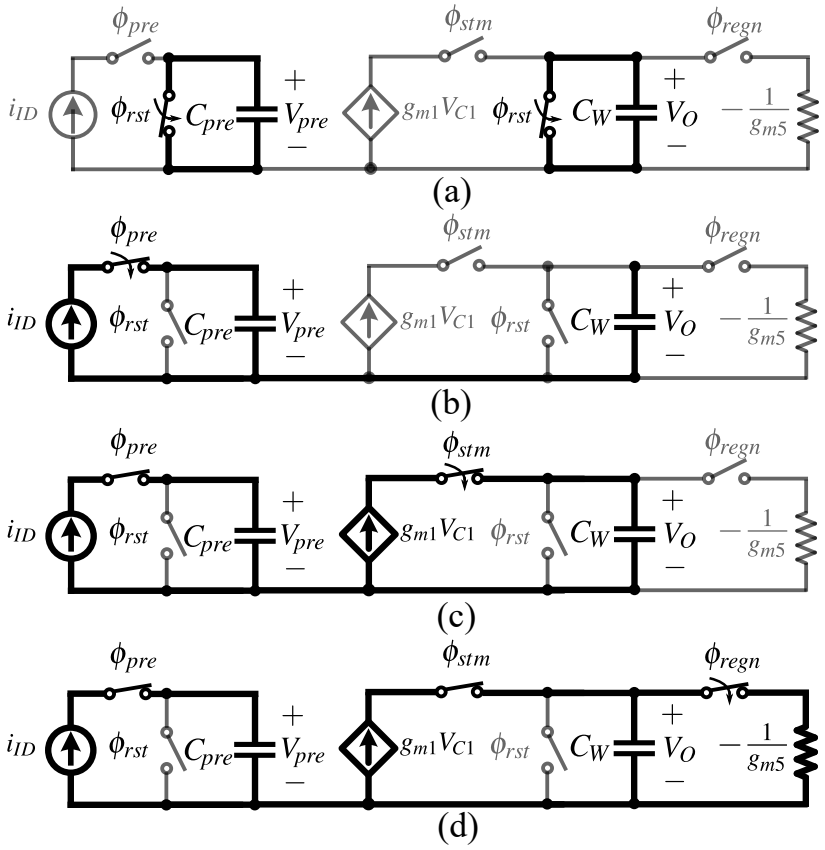


Figure 2.26: Equivalent differential circuits of a double-tail comparator during (a) reset phase; (b) first integration from dynamic preamplifier; (c) double integration from the dynamic amplifier and the second stage strongARM latch; (d) regeneration phase.

referred offset and noise.

2.7.1 Details of Operation

[25] offers a better circuit (Fig. 2.24(b)), recognizable as a dynamic amplifier MP1,MP2 driving a StrongArm latch in tandem. Why does this not suffer the same fate? It has to do with the *direction* of the ramping bias that the complementary (folded) amplifier applies to the StrongArm latch. The preamplifier drives the input pair M1a,M2a of the StrongArm latch with a positive-going ramp, on which is superposed the continuously growing differential signal. This ramp turns on the input pair of the StrongArm latch and gradually releases the reset switches M1b,M2b,M1c,M2c. The input pair's transconductance rises with the

ramp because the pair is in saturation, while its differential input grows continuously. The StrongArm goes through its self-timed phases of sampling, propagation, and regeneration. Since everything operates as expected, the comparator's input-referred offset and noise will now be determined almost entirely by the differential pair of the preamplifier.

The dynamic preamplifier will work correctly over a large range of input bias, so long as the FETs MP1,MP2 in its input pair remain in saturation. However, once triggered, it will always drive itself into the off condition. Its ramping output voltage gradually releases the reset switches and activates the tail current of the StrongArm. In sampling phase, the StrongArm itself will dynamically amplify the differential input further on C'_C. But if the ramping gate voltage of M1a,M2a becomes too large (or the supply voltage is low), they can discharge the amplified voltage stored on C'_C. As we have shown previously, the StrongArm will now display a large offset and noise. The preamplification is of little value if M1a and M2a, which are now the coupling FETs, enter triode and effectively disconnect the preamplifier before the StrongArm latch can regenerate.

This reveals that there exists within this cascade of two dynamic amplifiers [32, Sec. 5.2] a race condition that can undermine offset and noise. The dynamic preamplifier offers gain over a certain time window while its FET pair MP1,MP2 is in saturation. But when this FET pair enters triode, it effectively wipes out the amplified voltage stored on C_{pre}. The window is

$$T_{w,pre} = C_{pre}\frac{V_{IC} + |V_{t0}|}{\frac{1}{2}I_{pre}} \tag{2.52}$$

where, in Fig. 2.24(b), the voltages are measured with respect to ground. The StrongArm has its own window, $T_s + T_p$ determined by a changing bias current and the capacitances C'_C and C'_L.

The amplifier cascade will work well as long as $T_{w,pre} > (T_s + T_p)$. But if this inequality is reversed, the preamp will have erased its amplified voltage before the StrongArm starts to regenerate, resulting in a significant rise in input-referred offset. In reality the erasure is, of course, gradual. This is seen in simulations of the input-referred offset (Fig.2.27), with the offsets due to MP1,MP2 forming the baseline for comparison. The race condition, when

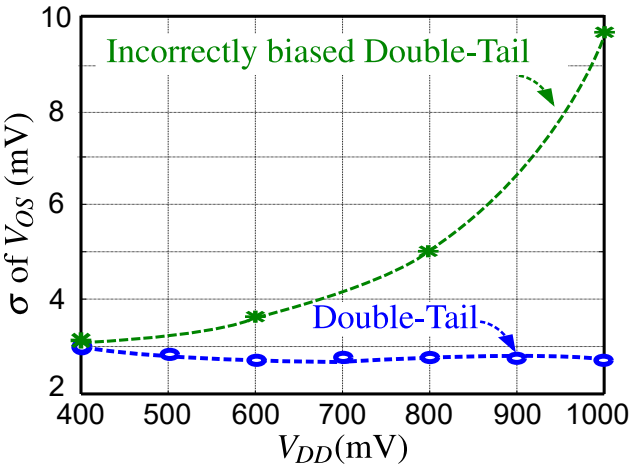


Figure 2.27: Offset degradation when M_B is removed and C_{pre} is reduced in Fig.2.24(b).

not accounted for in design by proper choice of C_{pre}/I_{pre}, is clearly damaging to the desired operation.

Comparator offset can be no lower than the input-referred offset of its input stage. The simplest method to calibrate this offset is with a digitally controlled array of capacitors inserted at the output of the dynamic preamp, as explained in Sec. 2.3. When realized correctly, this compensation will not be affected by changes in the supply voltage.

2.7.2 Noise in Double Tail

The classic StrongArm or the double tail latch will likely display comparable (uncompensated) *offset*, limited by mismatch in the preamplifying input pair; but one might believe that *noise* in the double tail should be lower because of the cascaded integration: that is, if windowed integrator bandlimits white noise, then a cascade of two such integrators should limit it even more. Analysis leads to the surprising result that this is not so: all else being the same, the noise in a double tail comparator is higher than it is in the StrongArm alone.

Suppose that the output of one windowed integrator drives, in tandem, another windowed integrator. After the first integrator has been active for a period T_{pre}, the second integrator is enabled with zero initial condition and driven by the output of the first; meanwhile the first continues to integrate (Fig. 2.26). The output of this second integrator is sampled after

an interval T_s. Then it can be shown (Appendix 2.14) that the net effect is that of sampling a white noise current bandlimited to $1/(T_{pre} + \frac{1}{2}T_s)$ Hz. This means that over a given time window $T_w = (T_{pre}+T_s)$, a single integration bandlimits the input noise more effectively than two integrations in tandem, where the second integration takes place over the sub-window T_s.

This calculation assumes time-invariant integrators, whose time constants are fixed. For proper autonomous (self-timed) operation in a double-tail (Fig. 2.24(b)), the sources of the input FET pair of the StrongArm latch should be grounded. This means that as the preamp's output bias voltage ramps up, the current through the StrongArm input stage rises quadratically with time and thus its transconductance linearly (Fig. 2.25). Therefore, its integration time constant also grows linearly with time. Since this time-varying circuit remains linear for small signals such as noise and offset, its time-dependent impulse response [48, p. 79] is a useful tool for analysis. A complete analysis (Appendix 2.15) shows that the input-referred mean square noise is set by the preamplifier's constant input spectral density (either voltage or current) limited to a NBW of $1/\left(T_{pre} + {}^{2}\!/_{3}(T_s + T_p)\right)$ Hz.

$$\langle v_{ID}^2 \rangle = \frac{8\mathrm{kT}\gamma}{g_{mp1,2}} \cdot \frac{1}{T_{pre} + \frac{2}{3}(T_s + T_p)} \tag{2.53}$$

Prior to the onset of regeneration, the total latency of the double tail is $T_{pre} + (T_s + T_p)$. If a simple StrongArm is designed for the same latency, it will bandlimit the input pair's noise to $1/(T_s + T_p)$ Hz, numerically equal to the inverse of the latency above but without the factor of $\frac{2}{3}$ in (2.53); its noise, therefore, will be slightly lower. This is because in the StrongArm, the charge integrated over T_s is transferred from C_C to C_L, so the process resembles a single, uninterrupted integration. On the other hand in a double tail cascade, the charge integrated by the preamplifier is *not* transferred to the integrator in the cascaded stage.

A good comparator design might trade off the slightly higher noise in the double tail against its robustness at low supply voltages or large spans of input common-mode levels. In view of the analysis so far, it is difficult to see what advantage in noise and offset that a third integrator in the cascade might offer, such as the triple tail described in [49].

2.8 Discussion: StrongArm vs. Double Tail

For the optimum trade-off among noise, offset and speed in the StrongArm latch, the input common mode must lie in a specified voltage interval. If too high, the input differential pair enters triode region prematurely, erasing the amplified input voltage during propagation phase; if too low, it stretches out latency and regeneration both. In practical designs, the input common mode is set by the output of the previous circuit and may not lie in the voltage range of preferred bias. For example, in certain successive approximation ADCs the input common-mode level changes at every conversion step [50]. There the double-tail comparator is more suitable [51] because its dynamic amplifier input can operate well across a wider range of input common mode. Ultimately this range is limited by the race condition described above.

The expected noise filtering can only be obtained with careful design. The first stage should provide sufficient amplification so that noise currents from M1a-c, M2a-c (Fig.2.24(b)) do not contribute significantly. Switches M1b,M1c when softly actuated by a ramp will inject noise of their own. The bias to the second stage latch is time-varying: M1a-c,M2a-c must be correctly sized to avoid the race condition—this may be the most important design step, and it requires proper choice of the constant current supplied by MB and of C_{pre} [52]. On the other hand, FET sizing is much easier in the classic StrongArm latch: a simple 1:1:1 size ratio of input differential pair, cross-coupled NMOS and PMOS pair is usually good enough.

Fig.2.28 and Fig.2.29 compare the input-referred noise, power consumption and delay of a StrongArm latch and a double-tail comparator. The two comparators are sized to give the same delay (at 1 mV differential input) and consume equal power at nominal 0.9V supply with $V_{in,CM} = 300\text{mV}$ in the StrongArm latch and $V_{IC} = 100\,\text{mV}$ in the double-tail comparator. As confirmation of its smaller NBW, the StrongArm latch displays 15% lower input-referred noise than a double-tail comparator; that is, until V_{DD} falls below 550 mV. At that point the input differential pair of the StrongArm latch is pushed into triode during propagation, resulting in erasure of dynamic amplification and rapid rise in noise. Although the double-tail's input-referred noise is somewhat higher, it is supply-independent. At $V_{IC} = 400\,\text{mV}$,

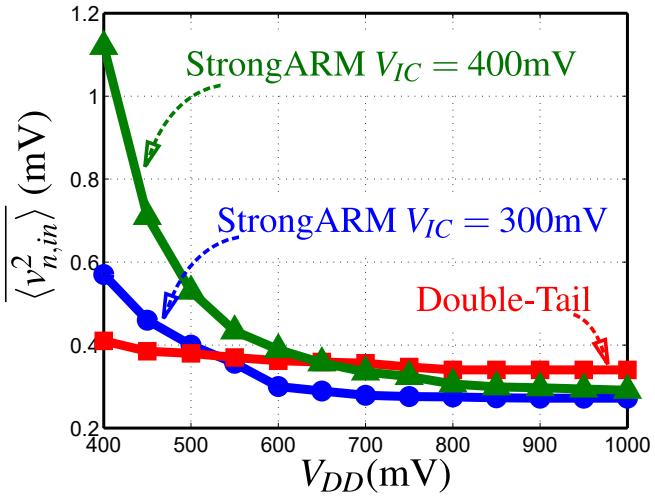


Figure 2.28: Input referred RMS noise of a strongARM latch and a double-tail comparator under different power supply levels.

noise in the StrongArm latch worsens even earlier at $V_{DD} = 700\,\text{mV}$, when the double-tail comparator offers lower noise.

The design trade-off is clear. We have proved that the StrongArm latch is the most compact yet efficient form of a regenerative comparator. Conceptually, it is a single-windowed integrator and a regenerative unit. Departure from this structure, such as extra integration or cascading multiple stages of dynamic amplifiers, tends to worsen noise. When the bias conditions allow, the StrongArm latch is to be preferred. But at low supplies or when the input biasing conditions are not suitable, the double-tail comparator is better. Its folded complementary preamplifier offers a useful level shift.

The double tail also offers a shorter regeneration time. We have described above how in the StrongArm regeneration starts with the transconductance of the common source pair M5,M6. As regeneration proceeds, the cross-connected pair M3,M4 gradually adds its own transconductance, which however is limited by the series resistance of M1,M2 in triode. Whereas in the classic StrongArm this resistance is determined by the fixed gate bias at M1,M2, in the double tail that bias voltage is ramping up which continuously lowers the degeneration resistance. This raises the time-dependent effective G_m and speeds up regeneration. In practice, this can lower regeneration time by $\sim$20%.

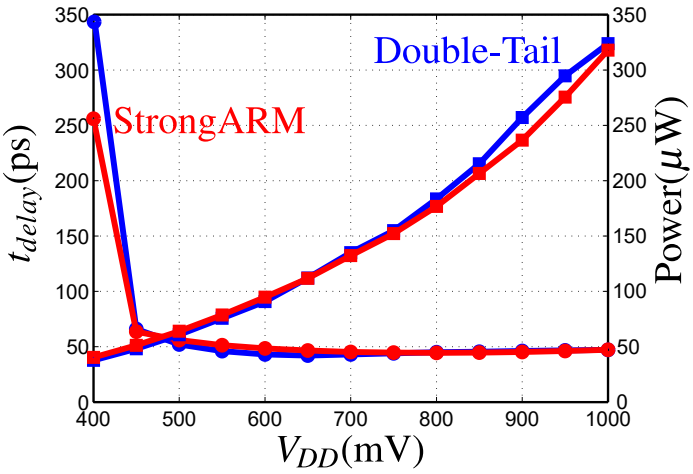


Figure 2.29: Power and delay of a strongARM latch and a double-tail comparator under different power supply levels. Clock frequency f_{CLK} = 2GHz and delay t_{delay} is simulated with 2mV differential input.

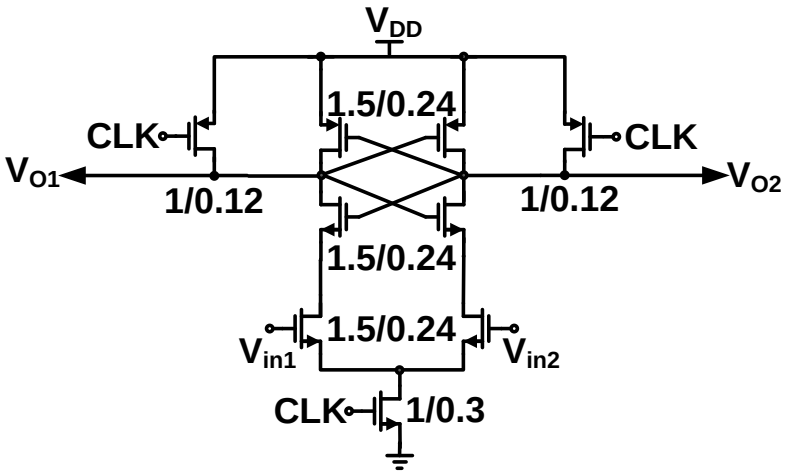


Figure 2.30: Test circuit in [1]. All dimensions in μm, V_{DD}=1.5 V.

2.9 Comparison with Measured Data

2.9.1 Input referred offsets of StrongARM Latch

All the expressions relating to static and dynamic offsets closely match the offset derived from transient simulations of the StrongARM latch as its input is swept. On the other hand, it is surprisingly difficult to find measured data in the many publications on this comparator in a

	StrongARM	Double-tail
Latency	$T_s + T_p = \frac{(C_L + C_C)V_t}{I_0}$	$T_{pre} + T'_s + T'_p = \frac{C_{pre}V_t}{I_0} + \frac{(C'_L + 2C'_C)V_t}{I'_0}$
NBW	$\frac{1}{2(T_s + T_p)}$	$\frac{1}{2[T_{pre} + \frac{2}{3}(T'_s + T'_p)]}$
Input referred noise	$\langle v^2_{n,stm} \rangle = 4\text{kT}\gamma \cdot \frac{1}{g_{m1,2}} \cdot \frac{1}{T_s + T_p}$	$\langle v^2_{n,D-T} \rangle = 4\text{kT}\gamma \cdot \frac{1}{g_{mp1,2}} \cdot \frac{1}{T_{pre} + \frac{2}{3}(T'_s + T'_p)}$
Power	$f_{CLK}(C_C + C_L)V^2_{DD}$	$f_{CLK}(C_{pre} + C'_C + C'_L)V^2_{DD}$
Scaling supply	Limited	Yes
Changing input common mode	Limited	Yes

Table 2.1: Comparison between a strongARM latch and a double-tail comparator.

	$V_{IC} = 1.05V$	$V_{IC} = 1.5V$
Measured Offset	8.5 mV	19.0 mV
Calculated Offset	9.5 mV	19.1 mV
Due to $\Delta V_{t1,2}$	5.8 mV	5.8 mV
Due to $\Delta \beta_{1,2}$	7.4 mV	14.0 mV
Due to $\Delta V_{t3,4}$	2.0 mV	11.6 mV

Table 2.2: Measured RMS offset vs. calculated. FET mismatch parameters: $A_{Vt} = 3.5\text{mV}{\cdot}\mu\text{m}$, $A_\beta = 2.5\%{\cdot}\mu\text{m}$.

form that could be used to validate the analysis developed in this paper: [1] is the exception, because it provides histograms measured across 45 samples of a latch realized in 130-nm CMOS. Notably the measured offset grows by more than 2× when the input common-mode exceeds the limit specified by (2.32). The circuit in question is shown in Fig. 2.30.

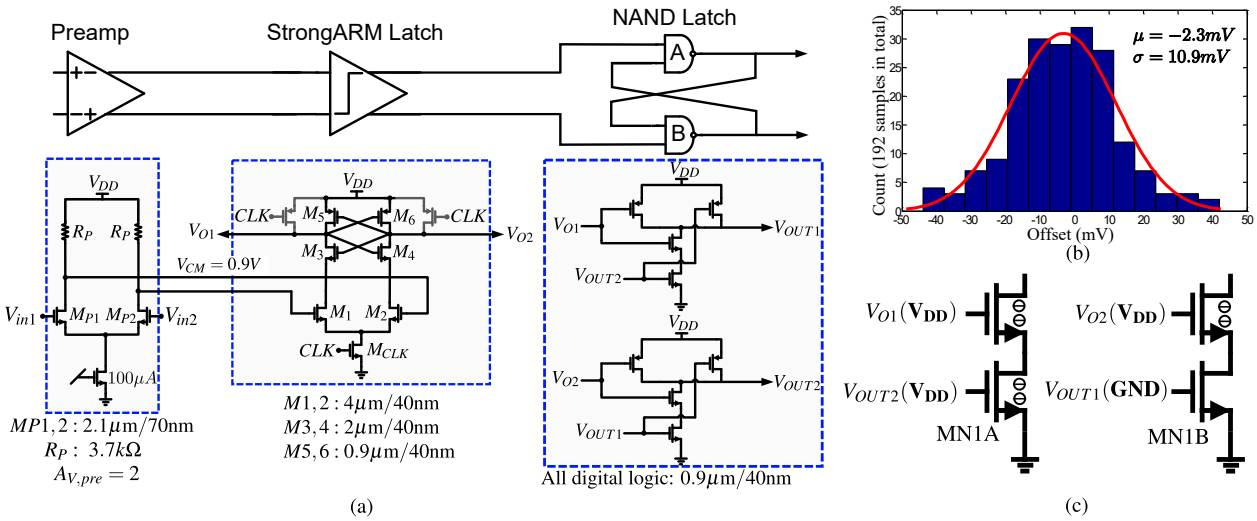


Figure 2.31: (a) A strongARM latch preceded by a differential preamp and loaded by a cross-coupled NAND latch, $V_{DD} = 1.05V$; (b) Measured input referred offset distribution; (c) Unbalanced capacitive load to strongARM latch from cross-coupled NAND gates.

Table 2.2 shows that the measured RMS offset at two common-mode levels agrees very well with predictions from our analysis. This is more than a matter of making the numbers come out close; our predictions were made to verify certain hypotheses. Although Infineon fabricated the circuit being measured, we use publicly available mismatch coefficients taken from the TSMC 130 nm process.

For $V_{IC} = 0.7\,V_{DD}$ which lies within the range in (2.32), we assume that offset arises *only* from V_t mismatch and β mismatch in M1,M2. Since our prediction is very close (for this size of population) to the total measured offset, it verifies the hypothesis that due to internal amplification, M3-M6 will not contribute appreciably to mismatch. For $V_{IC} = V_{DD}$ the analysis is more complicated because the gain $V_{O1,2}/V_{ID}$ changes over the two sub-phases described in Section 2.4.6. The detailed calculation is in Section 2.13.1 including the offset from V_t mismatch in M3,M4 (β mismatch doesn't matter because these FETs are in series with M1,M2). The prediction in this case is accurate. In spite of the lower internal amplification, mismatch in M5,M6 still does not contribute. However, it is clear that too large a V_{IC} will worsen the comparator's offset considerably.

Mismatch source	Original	Improved
Systematic offset due to hysteresis	±3.6mV	**±0.7mV**
Voltage gain of preamp $A_{V,pre}$	2	**4⇑**
Internal gain of the strongARM latch $A_{V,latch}$	1.5	**10⇑**
$\Delta\beta$ of diff. pair in preamp	6.5mV	**4mV**
ΔV_t of diff. pair in preamp	5.7mV	**4mV**
ΔV_t of M3,M4 in strongARM latch	3.2mV	**0.3mV**
ΔV_t in diff. pair in strongARM latch	2.8mV	**1.4mV**
$\Delta\beta$ in diff. pair in strongARM latch	1.4mV	**0.7mV**
Total random input-referred offset	10mV	**6mV⇓**

Table 2.3: Measured RMS offset vs. calculated. FET mismatch parameters: A_{Vt} = 2.3mV·μm, $A_\beta = 1.5\%\cdot\mu$m.

2.9.2 StrongARM Latch with Preceding Static Preamplifier and Loading NAND Latch

Shown in Fig.2.31(a) is a fabricated comparator from a commercial product in 90nm CMOS technology under 1.05V supply. It consists of three blocks, a preamplifier biased with a constant tail current source, a clocked strongARM latch and two cross-coupled NAND gates. The preamplifier preceding the strongARM latch shields the input differential pair from the strongARM latch and its associated dynamic offset due to clocking action. The preamplifier is meant to provide a voltage gain so that the strongARM latch does not contribute offset or noise to the input of preamplifier. The strongARM latch has the same configuration as in Fig.2.15. The cross-coupled NAND latch speeds up the regeneration and provides a stable output that remains steady while the strongARM latch resets. Thus, the only dominating offset source was supposed to be the mismatch between MP1,MP2.

However, the actual offset measurement in Fig.2.31(b) shows two unexpected problems. First, the mean of the measured offsets is a non-zero value, suggesting that this comparator has a systematic offset. Second, σ of the measured offset is much larger than expected, implying significant mismatch contributions other than MP1, MP2. Although the designers had difficulties explaining these phenomena, our analysis can accurately calculate the measurements, isolate sources of mismatch and furthermore, provide remedies to reduce offset without incurring penalty of extra power.

Each output of the strongARM latch is loaded by two NMOS transistors and two PMOS transistors from the NAND latch. Since the NAND latch stores '0' and '1' at its outputs that do not reset with the clock, when V_{O1},V_{O2} are reset to V_{DD}, one of the NMOS transistors (MN1A driven by $V_{OUT1}(V_{DD})$ for example) will be in strong inversion whereas the other (MN1B driven by $V_{OUT2}(GND)$ for example) in cut-off region as shown in Fig.2.31(c). Because of the extra electrons from the inverse layer in MN1A, V_{O1} experiences slower discharging than V_{O2}. This is equivalent to unbalanced capacitive loading ΔC_L to the strongARM latch, which further introduces a dynamic offset also known as the *hysteresis* effect. Under $V_{DD} = 1.05$V and $V_{in,CM} = 900$mV, output of the preamplifier has a common mode voltage of 900mV. This high common mode voltage pushes M1,M2 in the strongARM latch into triode during propagation phase. It degrades the internal voltage amplification within the strongARM latch to only about 1.5. With a voltage gain of 2 from the preamplifier, the whole comparator is calculated to have an offset hysteresis of $\pm$3.6mV, depending on the state of previous comparison cycle. In measurement, because the comparator is driven by a ramping voltage of the same slope to measure its trip point, the repeating unbalanced capacitive loading creates a systematic negative offset at the comparator's input.

Except threshold mismatch and β mismatch between MP1,MP2, threshold mismatch between M3,M4 also contributes significantly to the comparator input. Again, this is caused by the high common input voltage to the strongARM latch. Calculated breakdown of mismatch sources is shown in Table.2.3. The calculated net input referred offset matches well with measurement.

Our analysis provides a simple remedy to this comparator circuit without altering the

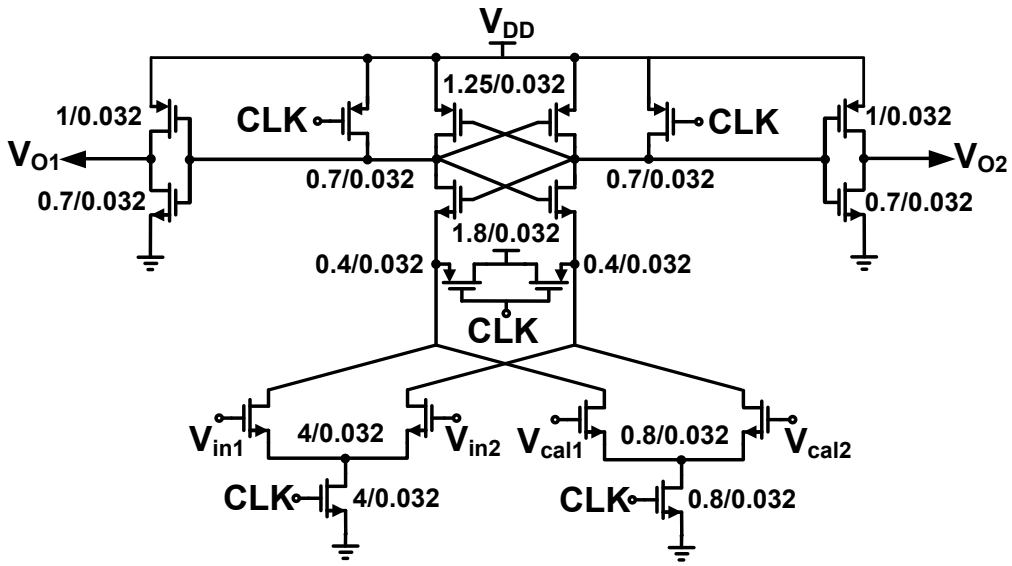


Figure 2.32: Test circuit in [2]. All dimensions in μm, V_{DD}=1 V.

circuitry or consuming more power. First, increasing sizes of M1,M2 by 2 times directly reduces the most dominating sources of mismatch. The second, yet more importantly, is to increase R_P in the preamplifier from 3kΩ to 8kΩ. This not only increases the preamplifier gain, but also reduces the preamplifier's output common mode voltage from 900mV to 650mV. The proper common mode input voltage to the strongARM latch raises its internal voltage amplification from 1.5 to 10, suppressing the mismatch contribution from M3,M4 to a negligible level. The raised voltage gain also shields the strongARM latch from the hysteresis effect. The predicted improvement with this simple remedy is shown in Table.2.3.

2.9.3 Input referred noise of StrongARM Latch

The input referred RMS noise of a strongARM latch is simulated versus the input common mode voltage in [2]. Using (2.50) predicts the simulated input referred RMS noise accurately as in Fig.2.33. When the input common mode voltage exceeds 600mV, the predicted results deviate from the simulated results, showing a smaller input referred RMS noise. This is because noises from M3,M4 begin to contribute when the high input common mode voltage pushes M1,M2 into triode. Calculation including impact from M3,M4 is shown in

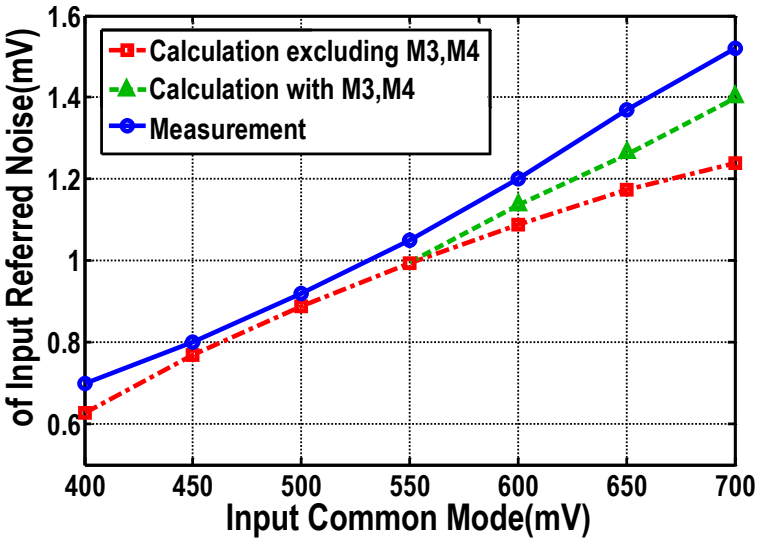


Figure 2.33: Calculated input referred noise versus measurement results in [2], parameter used in calculation: $\gamma = 2/3$, $T = 300°$, $V_T = 300\text{mV}$, $V_{P1,2} = V_{IC} - V_T$; $C_L = 2.6\text{fF}$ and $C_C = 1.1\text{fF}$ are estimated using $t_{OX} = 1.5\text{nm}$, $C_{OX} = 23\text{fF}/\mu\text{m}^2$.

Section2.13.2. Again, this underscores the importance of choosing the correct common mode voltage to achieve a high internal amplification.

2.10 Conclusion

We have developed a simple, physically-based analysis for the internal workings of a regenerative comparator. It recognizes the symmetry of the circuit, by identifying common mode and differential mode. The large-signal behaviour is readily understood by plotting equilibria and trajectories in the phase plane *defined by these modes*. It is most fruitful to consider comparator offset and noise as perturbations on the separatrix, the phase plane trajectory that leads from the initial reset condition into the metastable equilibrium. This brings clarity to long-debated distinctions between static and dynamic offsets in regenerative comparators. Simple equivalent circuits are shown to capture all time-varying aspects. For small offsets, nonlinearity in FET g_m does not change the shapes of the trajectories.p

These insights guide design of the widely used StrongARM comparator. Following reset,

three phases of operation are identified, with the role of pairs of FETs. An internal amplification is revealed, and from it follows the need to choose the input common-mode voltage correctly, for uses where the offset and noise should be kept very small. Most importantly, it is shown how by using a small array of switched capacitors, a dynamic offset can be made to cancel the static offset with little penalty on speed. This calibration is robust against supply fluctuations.

Predictions from the analysis are validated against measurements taken on a prototype StrongArm comparator. The measurements illustrate the major design considerations to emerge from the analysis. Our analysis of offset and its implications on design is much more direct than previous work on the topic, such as [21].

We are grateful to Professor Bernhard Wicht of Reutlingen University, Germany, who shared sizes of the FETs in the SRAM sense amplifier that he had reported in a perceptive early work.

2.11 Appendix

2.12 Output noise power in a strongARM latch

The complete expression of total output noise power at the end of propagation phase is the summation of (2.42) and (2.47), thus

$$\begin{aligned}\langle v_{n,out}^2 \rangle &= \langle v_{n,out1}^2 \rangle + \langle v_{n,out2}^2 \rangle \\ &= S_{in} \left[\frac{T_s + T_p - \tau_p}{(C_L - C_C)^2} - \frac{2(T_s - \tau_p)}{(C_L - C_C)^2} \cdot \mathrm{e}^{+s_p T_p} \right. \\ &\qquad \left. + \frac{T_s - \tau_p}{(C_L - C_C)^2} \cdot \mathrm{e}^{+2 s_p T_p} \right] \end{aligned} \tag{2.54}$$

With $C_C < C_L/2$, $T_p \gg \tau_p$ and $\exp(s_p T_p) \ll 1$, the total output noise power can be approximated as

$$\langle v_{n,out}^2 \rangle = S_{in} \cdot \frac{T_s + T_p}{(C_L - C_C)^2} \quad \text{if} \quad C_C < \frac{C_L}{2} \tag{2.55}$$

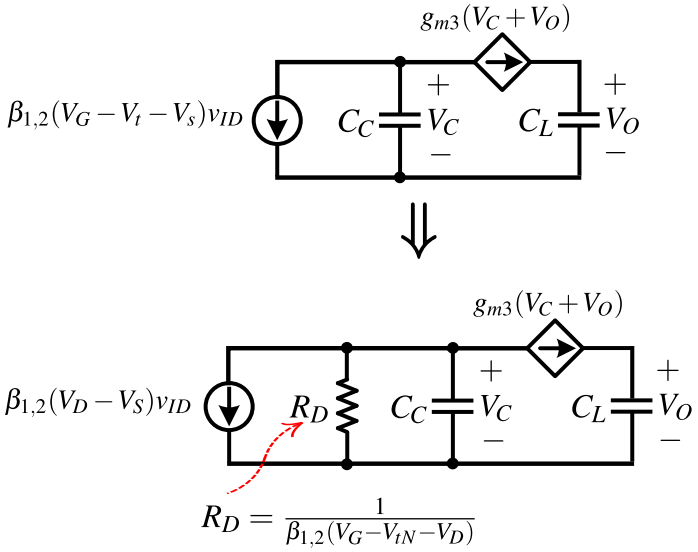


Figure 2.34: Equivalent differential circuit during propagation phase when $M1$ & 2 transit from saturation to triode region.

With $C_C \approx C_L$, $\tau \to \infty$ and $\exp(s_p T_p) \to 1$, using $e_x \approx 1 + x + x^2/2$, the total output noise power is approximated as

$$\langle v_{n,out}^2 \rangle = S_{in} \cdot \frac{T_s + \frac{1}{3}T_p}{(C_L - C_C)^2} \left(\frac{T_p}{\tau_p}\right)^2 \quad \text{if} \quad C_C \approx C_L \tag{2.56}$$

When $C_C > 2C_L$, $\exp(s_p T_p) \gg 1$ and the total output noise power is found by

$$\langle v_{n,out}^2 \rangle = S_{in} \cdot \frac{T_s - \tau_p}{(C_L - C_C)^2} \cdot \mathrm{e}^{+2 s_p T_p} \quad \text{if} \quad C_C > 2C_L \tag{2.57}$$

2.13 M1,M2 in triode during propagation phase

Instead of traditional square law device, here we choose EKV model to characterize the strongARM latch for its accurate and continuous representation of FET characteristics spanning saturation to triode regions. By decomposing the drain current into a forward current $I_{\mathscr{F}}$ that is independent of drain voltage, and reverse current $I_{\mathscr{R}}$ independent of source voltage.

$$I_{\mathscr{F}} = \frac{\beta}{2}(V_G - V_t - V_S)^2 \tag{2.57a}$$

$$I_{\mathscr{R}} = \frac{\beta}{2}(V_G - V_t - V_D)^2 \tag{2.57b}$$

The corresponding differential current is then also decomposed into a forward current and reverse current,

$$\Delta I_{\mathscr{F}} = \beta(V_G - V_t - V_S)\Delta v_G \tag{2.57c}$$

$$\Delta I_{\mathscr{R}} = \beta(V_G - V_t - V_D)\Delta v_G - \beta(V_G - V_t - V_D)\Delta v_D \tag{2.57d}$$

The net differential current between M1,M2 is calculated by subtracting $\Delta I_{\mathscr{R}}$ from $\Delta I_{\mathscr{F}}$

$$\begin{aligned}\Delta I_{M1,2} &= \Delta I_{\mathscr{F}} - \Delta I_{\mathscr{R}} \\ &= \underbrace{\beta_{1,2}(V_D - V_S)v_{ID}}_{\text{Diff. current}} + \underbrace{\beta(V_{IC} - V_t - V_D)V_{CD}}_{\text{CM coupled res.}}\end{aligned} \tag{2.58}$$

In the differential equivalent circuit during propagation phase as shown in Fig.2.34, this differential current can be further decomposed into two linear circuit elements, a constant current source i'_{ID} and a linear resistor R_D. However, It should be noticed that although i'_{ID} and R_D appear as linear circuit elements in the differential equivalent circuit, they are both cross coupled with the common voltage and their values change along with common mode voltage excursions.

The duration while M1,M2 remain in saturation, if there exists, is found by

$$\begin{aligned}T_{p,sat} &= \frac{(C_L + C_C)[(V_{DD} - V_t - V_{P3,4}) - (V_{IC} - V_t)]}{I_0} \\ &= \frac{(C_L + C_C)(V_{DD} - V_{IC} - V_{P3,4})}{I_0}\end{aligned} \tag{2.59}$$

Through this period of time, the circuit behaves the same as described in Sec.2.4.2.3, yet the equivalent integration window is shortened. Replacing T_r with $T_{r,sat}$ in (2.26), $A_{G,sat}$ and the preamplification voltage gain $A_{V,sat}$ can be then calculated.

When the drain voltages of M1,M2 become too low, M1,M2 are pushed into triode. The duration while M1,M2 remain in saturation before M5,M6 become active can be quantified as

$$T_{p,trd} = \frac{(C_L + C_C)(V_t - V_{DD} + V_{IC} + V_{P3,4})}{I_0} \tag{2.60}$$

The differential equivalent circuit in Fig.2.34(b) has two real poles, one in left half plane and the other in right half plane.

$$\begin{cases} s_{pl} = -\dfrac{1}{(R_D||g_{m3,4}^{-1})C_C} \\ s_{pr} = \dfrac{1}{(R_D + g_{m3,4}^{-1})C_L} \end{cases} \tag{2.61}$$

The output voltage is mainly dominated by the right half plane pole and can be approximated with two terms, first, the output voltage from integration during $T_{r,sat}$ will be regenerated by the right half plane pole $\mathbf{s}_{p2}$, second, i_{ID} flows through R_D, establishing a voltage across R_D that is also regenerated by $\mathbf{s}_{p2}$. Thus, the differential voltage on C_L for M5,M6 to finally regenerate in regeneration phase is the superposition of the two,

$$V'_{OD} = (v_{ID} \cdot A_{V,sat} + i'_{ID} \cdot R_D)\mathrm{e}^{+s_{pr}T_{p,trd}} \tag{2.62}$$

The voltage amplification gain before PMOS begin to conduct is calculated with $A'_V = V'_{OD}/v_{ID}$.

2.13.1 Offset due to mismatch between M3,M4 in [1]

With R_D present, the output voltage established by i'_{ID} is given by

$$\begin{aligned} V_{OD}(t) &= (i'_{ID} \cdot R_D)\mathrm{e}^{+s_{pr}T_{p,trd}} \\ &= v_{ID} \cdot \frac{V_D - V_S}{V_{IC} - V_t - V_D} \cdot \mathrm{e}^{+s_{pr}T_{p,trd}} \end{aligned} \tag{2.63}$$

Threshold mismatch between M3,M4 results in a differential current flowing through the controlled source, which is also regenerated by the right half plane pole

$$V_{OD}(t) = g_{m3,4}\Delta V_{t3,4}\left(\frac{1}{g_{m3,4}}\right)\mathrm{e}^{+s_{pr}T_{p,trd}} \tag{2.64}$$

The input referred offset due to $\Delta V_{t3,4}$ is then calculated by counter balancing the contribution from $\Delta V_{t3,4}$

$$V_{OS} = \Delta V_{t3,4} \cdot \frac{V_{IC} - V_t - V_D}{V_D - V_S} \tag{2.65}$$

It is readily shown by (2.65) how input referred offset increases with high input common mode voltage.

2.13.2 Extra noise from M3,M4 in [2]

Noise calculation is performed on time domain using the same method in [53]. The noise current from M3,M4 begins to integrate on C_L after M1,M2 are pushed into triode. The transfer function from noise current of M3,M4 to the output is

$$H(s) = \frac{g_{m3,4}^{-1}}{R_D + g_{m3,4}^{-1}} \cdot \frac{1}{C_L} \cdot \frac{1}{s - s_{pr}} \left(1 - \mathrm{e}^{-sT_{p,trd}}\right) \tag{2.66}$$

Its corresponding impulse response is

$$h(t) = \frac{g_{m3,4}^{-1}}{R_D + g_{m3,4}^{-1}} \cdot \frac{1}{C_L} \left(1 - \mathrm{e}^{+s_{pr}t}\right) [\mathbf{u}(t) - \mathbf{u}(t - T_{p,trd})] \tag{2.67}$$

Its contribution to the output noise power can be calculated

$$\begin{aligned} \langle v_{n,out}^2 \rangle &= S_{M3,4} \int_0^{T_{p,trd}} \left[\frac{g_{m3,4}^{-1}}{R_D + g_{m3,4}^{-1}} \frac{1}{C_L} \left(1 - \mathrm{e}^{+s_{pr}t}\right) \right]^2 \mathrm{d}t \\ &\approx S_{M3,4} \left(\frac{g_{m3,4}^{-1}}{R_D + g_{m3,4}^{-1}} \right)^2 \cdot \frac{\tau_{pr}}{C_L^2} \cdot \mathrm{e}^{+2s_{pr}T_{p,trd}} \\ &= S_{M3,4} \frac{g_{m3,4}^{-2}}{(R_D + g_{m3,4}^{-1})C_L} \cdot \mathrm{e}^{+2s_{pr}T_{p,trd}} \end{aligned} \tag{2.68}$$

The total input referred noise is the net result of both M1,M2 and M3,M4.

$$\langle v_{n,in}^2 \rangle \approx \frac{S_{M1,2} \cdot g_{m1,2}^{-2}}{T_s + T_{p,sat}} + \frac{S_{M3,4} \cdot g_{m3,4}^{-2}}{(R_D + g_{m3,4}^{-1})C_L} \cdot \frac{\mathrm{e}^{+2s_{pr}T_{p,trd}}}{A_V' 2} \tag{2.69}$$

Compared to the scenario where M1,M2 remain in saturation, the total input referred noise increase for two reasons, first, the integration time is reduced, second, M3,M4 begin to contribute significantly.

2.14 NBW of a cascade of two integrators

The impulse response of a single-stage windowed integrator Fig.2.35(a) is

$$h(t) = \frac{1}{C}[\mathbf{u}(t) - \mathbf{u}(t - T)] \tag{2.70}$$

Its input referred noise and NBW are then calculated as

$$\langle i_{ID}^2 \rangle = \frac{S_{in}}{2} \frac{\int_0^{+\infty} h^2(t)\mathrm{d}t}{\left(\int_0^{+\infty} h(t)\mathrm{d}t\right)^2} = \frac{S_{in}}{T} \Rightarrow \mathrm{NBW} = \frac{1}{T} \tag{2.71}$$

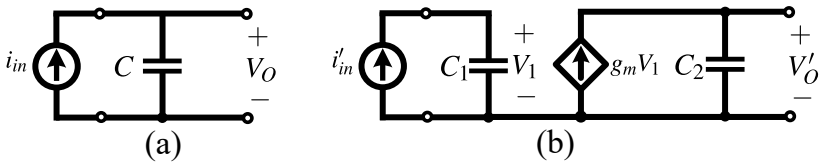


Figure 2.35: (a) Single integrator; (b) Cascading integrator.

In a two-stage cascading windowed integrator Fig.2.35(b), its impulse response is

$$h'(t) = \frac{1}{C_1} \cdot \frac{g_m t}{C_2}[\mathbf{u}(t) - \mathbf{u}(t-T)] \tag{2.72}$$

Its input referred noise and NBW are then calculated as

$$\langle i'^2_{ID} \rangle = \frac{S_{in}}{2} \frac{\int_0^{+\infty} h'^2(t)\mathrm{d}t}{\left(\int_0^{+\infty} h'(t)\mathrm{d}t\right)^2} = \frac{S_{in}}{\left(\frac{3}{4}T\right)} \Rightarrow \text{NBW} = \frac{4}{3} \cdot \frac{1}{T} \tag{2.73}$$

Indeed, cascading more integrators *widens* the NBW, which is undesirable. We offer a simple explanation. First, these are circuits with finite impulse response, where it is best to derive their frequency response with the Fourier transform. In a single windowed integrator, the rectangle impulse response transforms into the well-known sinc() function in frequency with nulls at the inverse of the integration time window and its multiples. But the impulse response of the integrator cascade is a single sawtooth triangle, whose more complicated transform [54, Fig. 10-16] is lowpass but with no nulls. It therefore transmits a larger mean-square noise.

2.15 LTV Noise Analysis of Double-tail Comparator

Impulse response $h(t, \tau)$ of a time-varying system describes the system response at time t to an impulse arriving at time τ. Since superposition still applies in a linear system, the output response

$$y(t) = \int_{-\infty}^{+\infty} h(t, \tau) \cdot x(\tau) d\tau \tag{2.74}$$

Output noise variance $\langle v^2_{n,out} \rangle$ sampled at time t_0 of the system is calculated by

$$\langle v^2_{n,out} \rangle = S_n \int_{-\infty}^{+\infty} h^2(t_0, \tau) d\tau \tag{2.75}$$

where $h(t_0, \tau)$ is the impulse response at time t_0 to an impulse arriving at time τ, S_n is power spectral density if the input.

Now we calculate the overall gain of the double-tail comparator. The differential voltage integrated on C_{pre} during the first integration is given by

$$V_{pre} = i_{ID} \times \frac{T_{pre}}{C_{pre}} \tag{2.76}$$

where T_{pre} is the duration of time when only the dynamic differential preamplifier integrates. After the second-stage becomes active, cascading integrators begin to integrate on C_W and the output differential voltage before regeneration is calculated by

$$\begin{aligned} V_{OD} &= \int_0^{T_W} g_{m2} \left(i_{ID} \frac{T_{pre}}{C_{pre}} + i_{ID} \frac{t}{C_{pre}} \right) \frac{1}{C_W} \, \mathrm{d}t \\ &= \frac{g_{m2} i_{ID}}{C_{pre} C_W} \left(T_{pre} + \frac{1}{2} T_W \right) T_W \end{aligned} \tag{2.77}$$

The transconductance gain of the double-tail comparator can be obtained with $A_G = V_{OD}/i_{ID}$.

Now we find the equivalent noise bandwidth of the double-tail comparator with the same method applied to the strongARM latch. From the first integration, the noise current from the input differential pair creates a noise voltage variance on C_{C1} of

$$\langle v_{n,p}^2 \rangle = S_{in} \int_0^{T_{pre}} \left(\frac{1}{C_{pre}} \right)^2 \mathrm{d}t = S_{in} \cdot \frac{T_{pre}}{C_{pre}^2} \tag{2.78}$$

The noise voltage on C_{C1} linearly scales to the output after the second stage becomes active,

$$\langle v_{n,out1}^2 \rangle = S_{in} \frac{T_{pre}}{C_{pre}^2} \left(\frac{g_{m2} T_W}{C_W} \right)^2 \tag{2.79}$$

Meanwhile, the noise current keeps integrating on C_W,

$$\langle v_{n,out2}^2 \rangle = S_{in} \int_0^{T_W} \left(\frac{1}{C_{pre}} \frac{g_{m2} t}{C_W} \right)^2 \mathrm{d}t \tag{2.80}$$

The total output noise voltage variance before regeneration is summation of the two contributions

$$\begin{aligned} \langle v_{n,out}^2 \rangle &= \langle v_{n,out1}^2 \rangle + \langle v_{n,out2}^2 \rangle \\ &= S_{in} \left(\frac{g_{m2}}{C_{pre} C_W} \right)^2 \left(T_{pre} + \frac{T_W}{3} \right) T_W^2 \end{aligned} \tag{2.81}$$

The input referred noise current variance can be then calculated

$$\langle i_{n,in}^2 \rangle = \frac{\langle v_{n,out}^2 \rangle}{A_G^2} = S_{in} \frac{T_{pre} + \frac{T_W}{3}}{\left(T_{pre} + \frac{T_W}{2} \right)^2} \tag{2.82}$$

The equivalent noise bandwidth NBW is now defined

$$\text{NBW} = \frac{T_{pre} + \frac{T_W}{3}}{\left(T_{pre} + \frac{T_W}{2} \right)^2} \approx \frac{1}{T_{pre} + \frac{T_W}{2}} \tag{2.83}$$

CHAPTER 3

Design Methodology for Phase-Locked Loops using Binary (Bang-Bang) Phase Detectors

We present a linearized analysis of bang-bang phase-locked loops (PLLs) in the frequency domain that is complete and self-consistent. It enables the manual design of frequency synthesis PLLs for loop bandwidth, output phase noise and minimum jitter. Tradeoffs between various parameters of the loop become clear. The analysis is validated against measurements on four very different loops, and helps to answer long-standing questions on aspects of these circuits attributable a hard nonlinearity. A brief designer's guide is included.

3.1 Introduction

The new generation of phase-locked loops (PLLs) on mixed-mode ICs is either partly digital [7] or all-digital [55, 56, 57, 58, 59, 11, 60, 61, 62], uses a time-to-digital converter (TDC) instead of a linear phase detector, and a digital filter in place of an the R-C filter network. The underlying concept dates back to compact spaceborne communications systems in the early 1970s; see, for instance, [63]. In this paper we examine partly or fully digital PLLs that use a bang-bang phase detector, a one-bit time-to-digital converter often realized by a regenerative clocked comparator as shown in Fig. 3.1[64, 62, 56, 11, 7]. It is similar to a one-bit delta-sigma A/D converter, which is preferred over a multibit converter when the conversion bandwidth allows. This is because a one-bit, two-level quantizer is perfectly "linear" in the sense that a straight line always passes through two points, whereas a multilevel quantizer can never be perfectly so because of practical sources of non-uniformity among more than two threshold levels.

A PLL using a bang-bang phase detector should be able to achieve a wider bandwidth than the TDC-based all-digital PLL, because a single comparator clocks at the maximum possible rate offered by a given IC technology. Unlike TDCs which generate digital words that can only be processed by digital filters, the '0' and '1' pulses from a bang-bang phase detector can be interpreted either as a digital bitstream or as a two-level periodically switched analog waveform. As an example of what is possible, [7] shows that despite the intrinsic quantization noise, the output spectrum of a bang-bang PLL with an analog loop filter can satisfy the stringent phase noise required of the GSM wireless receiver.

The design process of analog PLLs is mature, based on accurate analysis in the continuous-time phase domain (see, for example, [65, Ch. 9]). In spite of a growing body of literature on bang-bang PLLs, their design has not yet reached the same maturity. Often it will start with an accurate but complicated analysis, that quickly devolves into intensive trial-and-error simulations in the time-domain. What is needed is a method of analysis for this nonlinear feedback loop that enables a useful first-cut design based on manual calculations. Simulations are then relegated to functional verification and fine-tuning.

In this paper we will present a complete design-oriented analysis of the BB-PLL. We restrict ourselves to the case when the PLL input is periodic, that is, when the loop is being employed as a *frequency synthesizer*. We contribute the following beyond already published work:

1. An analysis entirely in the frequency domain;
2. That leads to familiar expressions for noise transfer functions (NTFs).
3. Calculation of the output phase noise as the superposition of the square magnitude of various NTFs.
4. Calculation of jitter using the noise bandwidth associated with each NTF. The dominant contributors to jitter become evident.
5. A proof in the frequency domain of why adaptation of the loop filter coefficients towards a certain autocorrelation of the binary sequence at the phase detector output leads to least jitter; and then an argument that adaptation may not be necessary.
6. A wide ranging validation against measurements reported in the literature on several

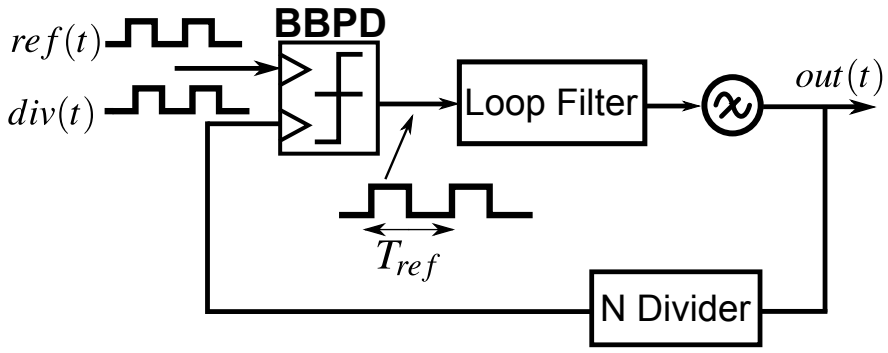


Figure 3.1: Phase-locked loop employing a binary bang-bang phase detector.

different BB-PLLs. This sheds light on various theoretical questions as they apply to practical circuits, such as if limit cycles will be seen, whether a Gaussian distribution adequately describes the jitter within the loop, and so on.

Two papers come close to certain aspects of what we will present, and we discuss them now. [66] derives the gain of the binary phase detector correctly, in line with our treatment. But it falls short of developing a closed-form expression for this gain in terms of loop parameters and noise levels. [67] goes further and provides closed-form expressions, but in terms of time-domain jitter quantities. We maintain that jitter is inconvenient as a principal working variable; instead it should be spectral density, the conventional form used to specify most noise sources. Jitter is then the product of known spectral densities with noise bandwidths.

Counter to the recent trend of analysis in the time-domain that seems to follow upon [68][69], we believe that design is a great deal simpler in the continuous frequency domain. It will resemble the now familiar design of analog PLLs where our intuition is well-developed. It holds up well until the loop bandwidth approaches half the reference frequency, when the granularity of phase detection in discrete-time (once every reference clock cycle) becomes prominent and the loop must be modelled more accurately in the z-domain [70].

Unlike a linear phase detector that generates a pulse width proportional to the time difference between its two input signals Fig.3.2(a), the binary phase detector is a one-bit time-to-digital quantizer with a discontinuous input/output characteristic, Fig.3.2(b). Its gain is not defined at zero input phase (the dot at the origin indicates a metastable state

[22]), and in all practical situations random noise will actuate the bang-bang phase detector around zero, which we will show enables an effective gain to be defined. Then from linear analysis, key quantities such as loop bandwidth are calculated.

An early work [64] models the bang-bang phase detector as a unit-gain element with quantization noise added to its output. The assumption of unity gain is simplistic: recent works turn to both nonlinear and linear analysis to deal with the phase detector's discontinuity. One approach [71, 72] linearizes the bang-bang phase detector by considering probability density functions, but the 'linearization' is valid only when the input signal is small [71]. Another characterizes the nonlinear loop with time-domain state equations but the title [73] notwithstanding, it is too complicated in our view to serve the needs of most circuit designers.

Time-domain and frequency-domain analysis are combined in [74], but the discussion there is restricted to limit cycles and spurs, excluding thermal noise in the loop. Other works first find the effective gain of the bang-bang phase detector by applying nonlinear time domain analysis [68] and then transform the bang-bang PLL to a linear system for phase noise analysis [75, 76, 77]. But as we will show, the effective gain of the bang-bang phase detector cannot be decoupled from phase noise. To calculate the phase detector gain, the total noise-induced time jitter at its input must be known in advance. This interdependence was clearly identified in a pioneering work on noise in feedback loops involving an element with a discontinuity [78].

The most accurate analysis published to date calculates the gain of the bang-bang phase detector by analogy to a Σ-Δ A/D converter [66]. An equivalent model has also been proposed in [72] based on the probability distribution function of the input phase error. But they fall short for practical purposes because neither develops a closed form analytical solution.

PLLs are used today for two distinct purposes: Frequency synthesizers, where output spectral purity is important, and clock and data recovery loops (CDR) where the objective is attenuation of input jitter.

In this paper we develop simple explicit expressions that depend only on known independent inputs to the loop and on values of design parameters, and we show how to use them to design a Frequency Synthesis loop to specifications. The design of BB-CDR loops involves other considerations because the inputs are random data waveforms, and special phase detectors are used whose gain depends on data statistics: space constraints oblige us into excluding their analysis here.

3.2 Effective Gain of the Bang-Bang Phase Detector

We start with a discussion of how to ascribe an effective gain to a phase detector (or for that matter any element) with a discontinuous characteristic. Once this gain is known and because all other elements are definable by transfer functions, linear system analysis can be applied to the entire bang-bang PLL. We must acknowledge that others have arrived at the same expression as ours for phase detector gain, but they do so either as an outcome of an *ad hoc* detailed analysis [68], or without connecting it clearly to the underlying mature body of knowledge [66].

Extracting the effective gain of a nonlinear device is a classic problem in system estimation. When a signal $x(t)$, either deterministic or random, is applied to a device with a nonlinear input/output mapping, the output $y(t)$ contains a component proportional to the input, and distortion. The bang-bang phase detector is an extreme example. Price's theorem [79] first proved that for a memoryless device, the cross-correlation between the input $x(t)$ and the output $y(t)$ will have the same shape as the autocorrelation of the input $x(t)$. An effective gain c may then be plausibly defined by the ratio between the two. This gain captures signal tranmission through the device. [80] discusses the error term that accounts for the distortion component arising from device nonlinearity. Thus, the bang-bang phase detector may be represented by a linear scaling gain block with an error added to its *output* [80],

$$y(t) = c \cdot x(t) + q(t) \tag{3.1}$$

where c is the linear scaling factor we call "gain" and $q(t)$ is the error waveform.

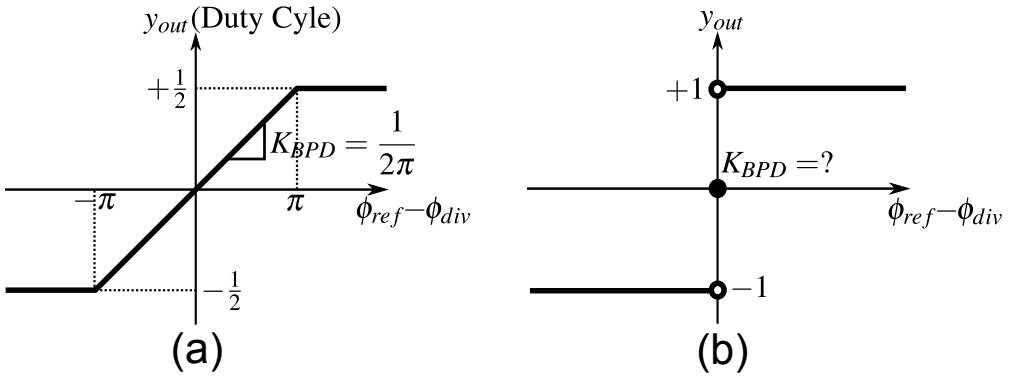


Figure 3.2: (a) Transfer function of a linear phase detector; (b) Transfer function of a bang-bang phase detector. Only the outputs $\pm\pi$ are stable.

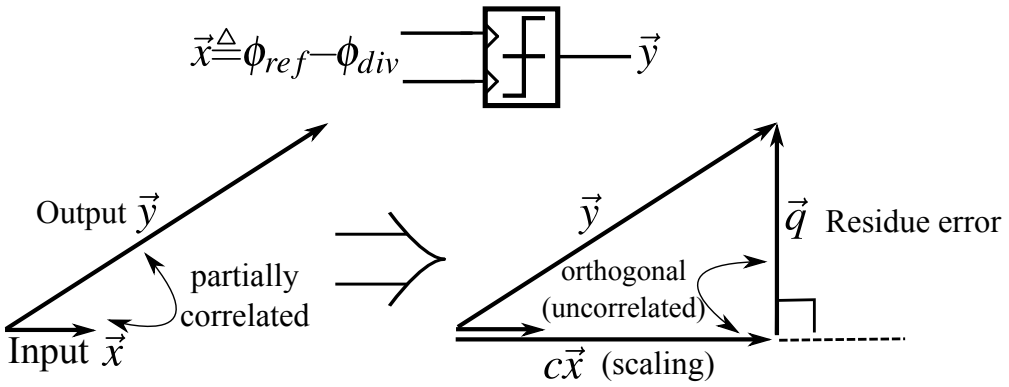


Figure 3.3: Interpretation of the gain extraction in *signal space*.

We will interpret this expression in signal space, where any signal is represented by a point in infinite dimensions that defines a vector from the origin [81, Sec. 6.10]. Since the output $\vec{y}$ is caused by the input $\vec{x}$, they must be correlated. But the discontinuity of the bang-bang phase detector distorts the output $\vec{y}$, forcing this correlation to be only partial. Estimation theory teaches that the effective gain is the c that minimizes the distortion or the residue error $\vec{q}$. As shown in Fig.3.3, this c is defined when the residue error $\vec{q}$ is orthogonal in signal space to $c\vec{x}$; since any other condition will result in a larger residue error $\vec{q}$. Thus, by multiplying both sides of (3.1) by $\vec{x}$ the effective gain c is the ratio of a cross-correlation

to an auto-correlation [81, Sec. 6.10][82, Ch. 4&5]

$$c = \frac{\vec{y} \cdot \vec{x}}{|\vec{x}|^2} \tag{3.2}$$

and the mean square of the error $q(t)$ is obviously

$$|\vec{q}|^2 = |\vec{y}|^2 - |c\vec{x}|^2 \tag{3.3}$$

With c as in (3.2), $\vec{q}$ becomes orthogonal to $c\vec{x}$, meaning $cx(t)$ and $q(t)$ are *uncorrelated*. This, incidentally, is the opposite of [76][77] which imply that $q(t)$ and $x(t)$ must be correlated because that produces a better fit to a certain curve.

Now we apply the concept of effective gain, $K_{BPD} \triangleq c$, to the bang-bang phase detector. In practical situations because there are many linear elements with memory present in the loop, the central limit theorem [83, Sec. 7.3] suggests that the input ϕ_e to the bang-bang phase detector, which is the phase difference between the reference clock and the divider output, has a Gaussian distribution with a zero mean and a standard deviation σ_{ϕ_e}. This assumption has been challenged [84] because the bang-bang loop is nonlinear. But in Section 3.7, by comparing the predictions of an analysis that assumes a Gaussian distribution against measurements and simulations, we will show that for a BB-PLL in the locked condition, this assumption holds up well for most practical purposes and greatly simplifies analysis. Perhaps this is so because this nonlinear element is memoryless.

The cross-correlation between the input and the output waveforms is calculated readily when input waveforms follow a Gaussian distribution. Since the output $y(t)$ is ± 1, $\mathsf{E}\{\phi_e(t)y(t)\}$ is the expected value of a full-wave rectified Gaussian waveform [54, Ch. 5].

$$\begin{aligned}
\mathsf{E}\{\phi_e(t)y(t)\} &= \int_{-\infty}^{0} (-1)\cdot\frac{\phi_e}{\sqrt{2\pi}\sigma_{\phi_e}} \exp\Big(-\frac{\phi_e^2}{2\sigma_{\phi_e}^2}\Big)\mathrm{d}\phi_e \\
&\quad + \int_{0}^{+\infty} (+1)\cdot\frac{\phi_e}{\sqrt{2\pi}\sigma_{\phi_e}} \exp\Big(-\frac{\phi_e^2}{2\sigma_{\phi_e}^2}\Big)\mathrm{d}\phi_e \\
&= 2\int_{0}^{+\infty} (+1)\cdot\frac{\phi_e}{\sqrt{2\pi}\sigma_{\phi_e}} \exp\Big(-\frac{\phi_e^2}{2\sigma_{\phi_e}^2}\Big)\mathrm{d}\phi_e \\
&= \sqrt{\frac{2}{\pi}}\cdot\sigma_{\phi_e}
\end{aligned} \tag{3.4}$$

Then the linear gain K_{BPD} of the bang-bang phase detector is

$$K_{BPD} = \frac{\mathsf{E}\{\phi_e(t)y(t)\}}{\mathsf{E}\{\phi_e^2(t)\}} = \sqrt{\frac{2}{\pi}} \cdot \frac{1}{\sigma_{\phi_e}} \tag{3.5}$$

From (3.3), the variance of the residue error signal $q(t)$ is

$$\sigma_q^2 = \mathsf{E}\{y^2(t)\} - \mathsf{E}\{[K_{BPD}\phi_e(t)]^2\} = 1 - \frac{2}{\pi} \tag{3.6}$$

The comparator toggles once per cycle at the reference frequency f_{ref}, capturing at that instant whether the PLL divider's rising edge lags or leads the transition of the reference clock. In other words, it *samples and holds* this binary phase comparison at a rate f_{ref}. For random phase changes caused by noise in one or both PD inputs, the phase error ϕ_e is confined to the interval $(-\pi, +\pi]$ because of the modulo-2π property of phase, and is uniformly distributed over it. The phase detector output toggles between ± 1, that is, it digitizes ϕ_e to 1 bit. Classic analysis of quantization noise (e.g. [83, Sec. 5.8]) tells us that the expected mean-square quantization error is $((+1) - (-1))^2 / 12 = 1/3$. This is very close in magnitude to the right-hand side of (3.6).

Quantization noise—which resembles wideband white noise of the same mean-square value—that is sampled-and-held at clock frequency f_{ref} defines a single-sided power spectral density on a continuous frequency axis of [85, Appendix B]

$$S_q(f) = \left(1 - \frac{2}{\pi}\right) \frac{2}{f_{ref}} \mathrm{sinc}^2\left(\frac{f}{f_{ref}}\right), \; \mathrm{sinc}(x) \triangleq \frac{\sin(\pi x)}{\pi x} \tag{3.7}$$

As $0.8 < \mathrm{sinc}^2(x) \leq 1$ for $0 \leq x < 0.25$, we can say that for most practical purposes $S_q(f)$ is constant over the frequency band $(0, f_{ref}/4)$, and may be modelled as white noise across it. In spite of the periodic clocking, this spectrum is continuous without discrete spectral lines.

This spectral density of the total phase detector output is zero at f_{ref}. Compare this with the output spectrum for a charge pump phase detector, which produces pulses whose width is proportional to the phase error sampled at a rate of f_{ref}. Whereas a bang-bang PD output is sampled-and-held, the charge pump's output current is effectively impulse sampled via pulse-width modulation. After conversion to a voltage in a loop filter which contains a series resistor, images of the baseband phase error spectrum will appear at f_{ref}

and its integer multiples [86]. Therefore, in frequency synthesizers using charge pumps where output spectral purity is at a premium, a second pole is often added to the loop filter to attenuate the image at f_{ref} with a second-order rolloff. By contrast, in BB-PLLs this second pole is unnecessary because of the intrinsic spectral null at f_{ref} and multiples. This explains the very low reference spurs (-72dBc/Hz) that have been reported, e.g. in [87, Fig. 16], for loops with binary phase detectors and a minimal loop filter of the form (3.8).

The method of signal decomposition described above is completely general. Although applied here to the bang-bang phase detector, it extends to all types of devices and to random or deterministic waveforms[1]. A multi-bit quantizer, also a nonlinear device, presents a well-defined gain with small RMS error q to an input signal that spans the whole input quantization range; the gain is defined by the best fit straight line through its input/ouput staircase characteristics: but when the input signal is so small that it toggles across only one quantization step, then like a bang-bang phase detector the phase detector gain changes with the RMS input (3.5).

(3.5) shows that an input signal ϕ_e with a larger RMS value will lower the effective gain of the bang-bang phase detector. This degenerated gain is undesired in most situations because it may result in higher in-band phase noise or smaller loop bandwidth. It can even underdamp the loop's transient response, because, depending on the order of the loop filter, a Type-II PLL may be conditionally stable; that is, the loop has a good phase margin only when the gain lies in a range with a well-defined minimum and a maximum. In the following sections, we use these ideas to explain findings reported in [75, 76, 9].

---

[1] A sinewave of amplitude $\sqrt{2}\sigma_{\phi e}$ at the input of the phase detector obeys a probability distribution that is quite the contrary to a Gaussian [54, Fig. 5-14]. Applying our analysis from first principles leads to $K_{BPD} = 2\sqrt{2}/(\pi\sigma_{\phi e})$. This will be recognized as the *describing function* of a step-like nonlinearity [88, p. 224][66], where now the error $q(t)$ is the sum of the 2nd and all higher harmonics in the output square wave. Since each harmonic is uncorrelated with every other one, so is their sum, satisfying the defining requirement for $\vec{q}$ in signal space.

3.3 Loop Noise Determines Gain of the Bang-Bang Phase Detector

To predict the phase locked loop dynamics, an effective gain K_{BPD} must be associated with the bang-bang phase detector. K_{BPD} is a function of the RMS input phase error σ_{ϕ_e} as in (3.5), but the various contributions to ϕ_e themselves depend on the loop dynamics (Fig.3.4). This results in an implicit relationship between the effective phase detector gain K_{BPD} and the phase error at its input σ_{ϕ_e}.

We will now obtain a simple closed-form expression for K_{BPD}.

3.3.1 Loop construction and model

A Type-II PLL is a cascade of two integrators in feedback, with a frequency compensation network inserted to reach the desired phase margin of stability. One of the integrations is implicit in the conversion of VCO output frequency to phase: it is this phase that actuates the phase detector. The other integrator may be realized in a variety of ways as part of the loop filter. For example, the simplest filter transfer function $F(s)$ used in a BB-PLL is the classic proportional-plus-integral type [89, Fig. 2.5]:

$$F(s) = \frac{\alpha f_{ref}}{s} + \beta \tag{3.8}$$

where α and β are dimensionless coefficients. This provides, in addition to an integration, a zero at the radian frequency $\omega_z = (\alpha/\beta) f_{ref}$.

Ascribing the gain K_{BPD} (as yet unknown in value) to the phase detector, Fig.3.4 shows the complete signal flow graph of the PLL. To avail the simplicity of transfer function analysis, all loop variables are expressed by their transforms in the Laplace domain. Since the phase detector output alternates between the dimensionless numbers ± 1 – obviously the quantization noise inherent in this output is also then dimensionless – it must be scaled by some voltage V_{FS} in order to drive the loop filter, and then the VCO. Often V_{FS} will be equal to the power supply voltage. K_{BPD} is of dimension rad^{-1}, and the VCO gain, K_V, is of dimension rad/(sec-V). In an all-digital realization with m-bit control, the product $K_V V_{FS}/2^m$

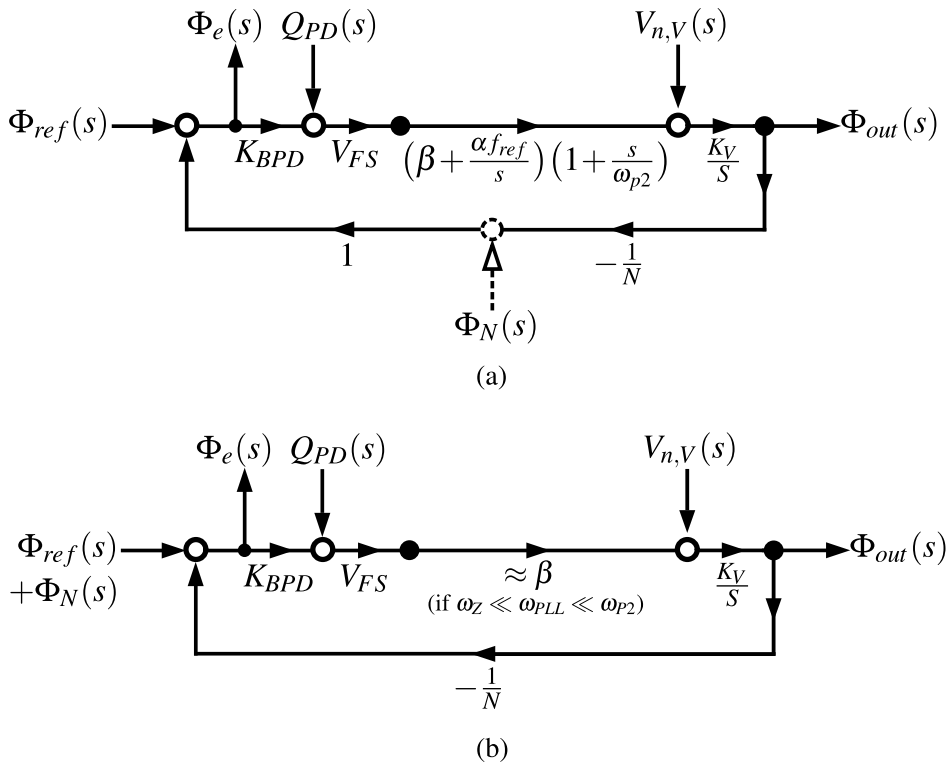


Figure 3.4: (a) Complete signal flowgraph of a PLL showing various sources of noise and where they are injected; (b) Simplified signal flow graph if $\omega_Z \ll \omega_{PLL} \ll \omega_{P2}$.

defines a VCO gain in units of rad/s/bit.

The loop gain $T(s)$ is simply the product of the transfer functions going once around the loop:

$$T(s) = K_{BPD} \cdot V_{FS} \cdot F(s)\frac{K_V}{s} \cdot \frac{1}{N} \tag{3.9}$$

The three sources of noise are shown as independent inputs injected into the signal flow graph at the appropriate nodes: phase noise on the periodic reference waveform ($\Phi_{ref}(s)$), quantization noise associated with the binary phase detector ($\mathcal{Q}_{PD}(s)$), and quantization noise associated with a possibly time-varying modulus of the frequency divider ($\Phi_N(s)$). Following [86], the VCO's own phase noise is modelled by an input-referred noise voltage $V_n(s)$ of constant power spectral density,

$$S_{Vn}(f) = K_W/K_V^2 \text{ V}^2/\text{Hz} \tag{3.10}$$

which produces at the VCO output the phase noise

$$S_{\phi_{VCO}}(f) = K_W/f^2 \text{ rad}^2/\text{Hz} \tag{3.11}$$

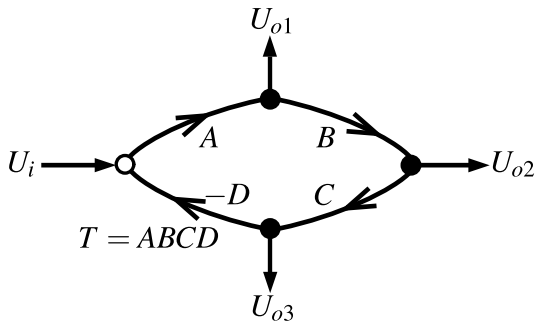
at the offset frequency f from the oscillation frequency f_0. This is consistent with a physically correct expression for the phase noise in an LC oscillator arising from white noise alone [8]. For a given LC quality factor, K_W depends only on power consumption in a certain technology, and on supply voltage [8]. When the VCO is a ring oscillator, (3.11) still describes the output phase noise but K_W is defined by a different set of frequency-independent circuit parameters [90].

Except for Φ_N, all three independent noise injections into the PLL's signal flowgraph are of constant spectral density, so their respective spectral contributions at the PLL output or at the phase detector's input will be simply a linear combination of the respective noise transfer functions. This simplifies the following analysis.

Unlike the charge-pump which captures detects both frequency and phase, a bang-bang phase detector can only detect phase difference, and requires a separate frequency capture loop. This has been referred to as a 'coarse tuning loop' [11, 7]. It prevents the PLL from locking at rational number multiples of the reference frequency. Because it is inactive after initial locking and does not interact with the bang-bang phase detector, aside from noting its necessity here it does not enter subsequent analysis.

3.3.2 Loop Transfer Functions

To minimize clutter in calculating loop transfer functions, we follow some guidelines in choosing between one of two possible formats. Fig.3.5 summarizes these alternatives for any transfer function in a feedback loop. Depending on the relative locations in the loop of the injection and of the response of interest, we will choose the form that uses the fewest parameters. Thus, the three sources of noise superpose as follows at the phase detector's



$$\frac{U_{o1}}{U_i} = \frac{A}{1+T} = \frac{1}{BCD} \times \frac{T}{1+T}$$
$$\frac{U_{o2}}{U_i} = \frac{AB}{1+T} = \frac{1}{CD} \times \frac{T}{1+T}$$
$$\frac{U_{o3}}{U_i} = \frac{ABC}{1+T} = \frac{1}{D} \times \frac{T}{1+T}$$

Figure 3.5: A feedback loop with one input and three possible outputs, and alternative formats of transfer functions.

input, or at the PLL's output (see Fig.3.4):

$$\Phi_e = \Phi_{ref} \frac{1}{1+T} \quad + \mathcal{Q}_{PD} \frac{1}{K_{BPD}} \frac{T}{1+T} + V_n \frac{K_V}{sN} \frac{1}{1+T} \tag{3.12}$$

$$\Phi_{out} = \Phi_{ref} N \frac{T}{1+T} + \mathcal{Q}_{PD} \frac{N}{K_{BPD}} \frac{T}{1+T} + V_n \frac{K_V}{s} \frac{1}{1+T} \tag{3.13}$$

Since the term $1 + T(s)$ appears in every denominator, its roots are the poles in the frequency response; more essentially $T(s)$ determines whether the loop is stable. To establish stability first, we discuss the qualitative features of $T(j\omega)$ that will lead to an adequate phase margin.

3.3.3 Loop Phase Margin

Fig.3.6(a) is an asymptotic Bode plot of $|T(j2\pi f)|$. Since T is always dimensionless, it is appropriate to use the unity gain frequency f_{PLL} as a reference value on the frequency axis. These features are salient:

1. For stable operation, a feedback loop with two integrators must always include a left-half plane zero at a frequency f_z such that $f_z < f_{PLL}$.
2. A filter pole might possibly be added at some $f_{P2} > f_{PLL}$ to more effectively attenuate the noise spectrum beyond f_{PLL}. This pole will tend to erode phase margin and the question is how to select its frequency. For greater than 60° phase margin, the net phase

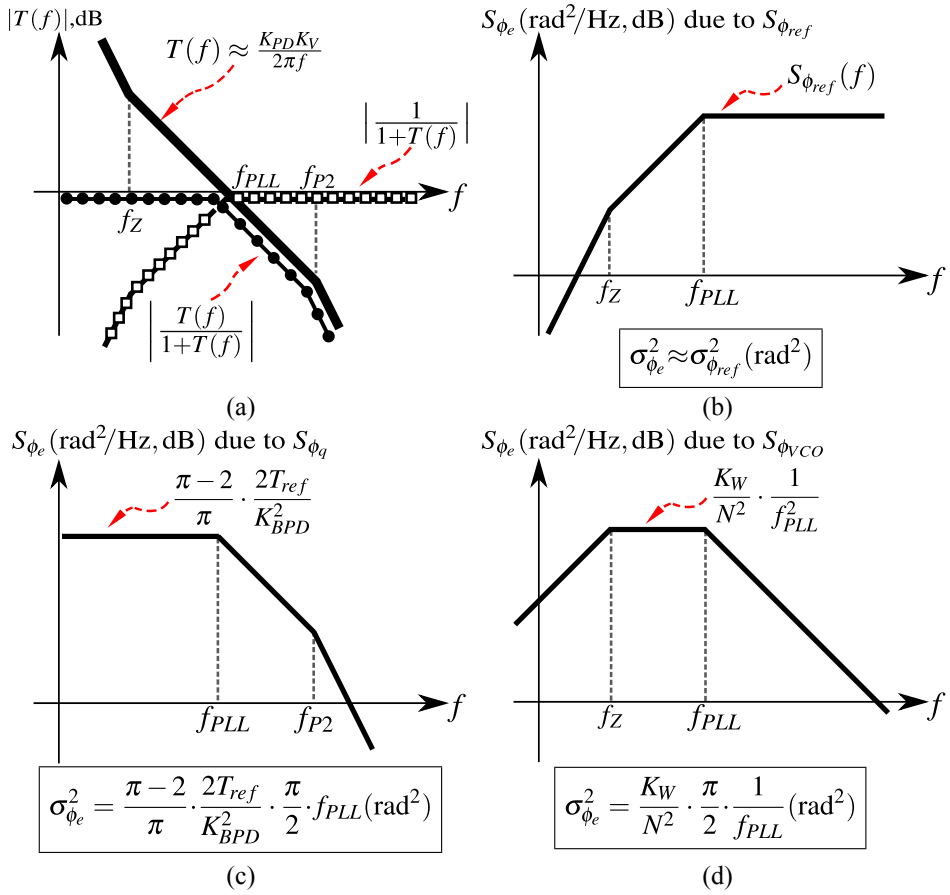


Figure 3.6: (a) Bode plot of loop gain magnitude, $|T(f)|$. Spectral density at PD input ($S_{\phi e}$) due to (b) reference noise; (c) PD quantization noise; (d) VCO phase noise.

lag due to the pole and zero at f_{PLL} must satisfy the inequality

$$\begin{aligned} &-\arctan\left(\frac{f_{PLL}}{f_{P2}}\right)-\arctan\left(\frac{f_Z}{f_{PLL}}\right)>-30^\circ \\ \Rightarrow\quad &\frac{\frac{f_{PLL}}{f_{P2}}+\frac{f_Z}{f_{PLL}}}{1-\frac{f_Z}{f_{P2}}}\simeq\frac{f_{PLL}}{f_{P2}}+\frac{f_Z}{f_{PLL}}<\frac{1}{\sqrt{3}} \end{aligned} \tag{3.14}$$

This inequality is readily satisfied when, in general, $f_Z \ll f_{PLL}$ and $f_{P2} \gg f_{PLL}$ because the zero gives a phase lead approaching 90° at f_{PLL} while the pole contributes a negligible phase lag.

Assuming that the loop is well-designed for a reasonable phase margin of at least 60°, we may deduce from (3.14) that the zero and pole frequencies each lie below and above f_{PLL} by a reasonable multiple such as 4×, thereby satisfying the inequality in (3.14) with a left-

hand side of roughly $\frac{1}{2}$. With these well-separated frequencies, we may write the following expression for $T(s)$ by looking at the piecewise asymptotes of the Bode plot Fig.3.6(a):

$$T(s) \simeq \frac{\omega_{PLL}}{s} \frac{1 + \frac{\omega_Z}{s}}{1 + \frac{s}{\omega_{P2}}} \tag{3.15}$$

When the filter has no second pole, we simply let $\omega_{P2} \to \infty$ in this expression.

3.3.4 Noise Bandwidths

In designing any PLL, one is called upon repeatedly to calculate the phase jitter or time jitter at the phase detector input or the PLL output. In the frequency domain, phase jitter is found by integrating phase noise spectral density over all frequencies. Since we have modelled all sources of noise thus far as white, the well-known concept of *noise bandwidth* simplifies the calculation of jitter with tidy analytical expressions. Although [89, Sec. 2.8] had advocated this approach for PLLs, it appears not to have been used so far.

All noise transfer functions for a Type-II PLL will be second-order lowpass $H_{LP}(s)$ or bandpass $H_{BP}(s)$. These transfer functions can always be expressed in one of the two standard formats:

$$H_{LP}(s) \;\; = H_0 \frac{1}{1 + \frac{s}{\omega_0 Q} + \frac{s^2}{\omega_0^2}} \tag{3.15a}$$

$$H_{BP}(s) = H_{max} \frac{\frac{s}{\omega_0 Q}}{1 + \frac{s}{\omega_0 Q} + \frac{s^2}{\omega_0^2}} \tag{3.15b}$$

where if $0 < Q < \frac{1}{2}$ the two poles are real, otherwise they are complex conjugate. ω_0 is always their geometric mean. After replacing s with $j\omega$, we use the integrals given in [91] to calculate the noise bandwidth (NBW) for the single-sided power spectral density. Thus,

$$NBW_{LP} \;\; = \omega_0 (Q/4) \text{ (Hz)} \tag{3.15c}$$

$$NBW_{BP} = \omega_0 \div (4Q) \text{ (Hz)} \tag{3.15d}$$

These expressions are exact, for any Q. The jitter is then simply the product of the constant spectral density of each injected noise, a scale factor, and the noise bandwidth (either (3.15c) or (3.15d)) associated with the applicable transfer function.

3.3.5 Jitter Calculations

We must know the jitter $\langle \phi_e^2 \rangle$ at the *input* of the binary phase detector to determine its gain. This jitter may be found using the transfer functions in (3.12) and the spectral density for the associated noise. First, we calculate the noise bandwidths. We determine if a coefficient on the right-hand side of (3.12) is lowpass or bandpass – or even highpass – and what are its ω_0 and Q. Assuming a robust phase margin, these functions can be evaluated with calculations directly on the graph. For example, the reference noise is transmitted through $|1/(1 + T(j\omega))|$, a highpass function with a lower corner frequency of f_{PLL} (Hz). But as a sampled-data system, the highest frequency it can transmit is $\frac{1}{2} f_{ref}$ (Hz) $= \pi f_{ref}$ (rad/s), which makes transmission effectively bandpass. For two widely separated poles the higher frequency pole is $\omega_0/Q = \pi f_{ref}$. Thus, using the expression (3.15d), in any second-order transfer function the noise bandwidth for the reference phase noise is $\omega_{ref}/8 = (\pi/4) f_{ref}$ (Hz).

Feedback remains effective up to the frequency of unity loop gain (f_{PLL}), when $|T(f_{PLL})| = 1$. As we saw in the calculation immediately above, this defines a corner frequency for all noise transfer functions in the loop. In a stable loop $f_Z \lesssim f_{PLL}/3$ while $f_{P2} \gtrsim 3 f_{PLL}$ as in Fig. 3.6(a). Then

$$f_{PLL} = \frac{\omega_{PLL}}{2\pi} \simeq \frac{\beta K_{BPD} V_{FS} K_V}{2\pi N} \tag{3.16}$$

The principal noise sources contributing to σ_{ϕ_e} are: (a) the VCO phase noise, (b) noise in the reference frequency source, and (c) the quantization noise arising from the bang-bang phase detector. Divider phase noise is usually not important, unless it is fractional-N quantization noise. It is also easily shown, given the short rise and fall times of regenerative comparators realized in state-of-the-art technologies and knowing their equivalent input RMS noise voltage [22], that a high-speed phase comparator contributes negligible jitter.

The phase noise voltage S_{vn} associated with the VCO experiences a bandpass transfer function to the input of the phase detector as in Fig.3.6(d). To see this, select the third term on the right-hand side of (3.12) then use the expression in (3.15), and ignore the filter pole

ω_{p2} to obtain

$$\begin{aligned}\Phi_e(s) &= V_n(s)\frac{K_v}{s\cdot N}\frac{1}{1+\frac{\omega_{PLL}}{s}(1+\frac{\omega_Z}{s})} \\ &= V_n(s)\frac{K_v}{N\omega_{PLL}}\frac{s/\omega_Z}{1+\frac{s}{\omega_Z}+\frac{s^2}{\omega_Z\omega_{PLL}}}\end{aligned} \tag{3.17}$$

This may be compared with (3.15b). For a loop with good phase margin $Q=\omega_Z/\omega_{PLL}$ is $\ll \frac{1}{2} \Rightarrow \omega_0/Q \simeq \omega_{PLL}$. Then referring to the expression for noise bandwidth (3.15d), $NBW_V = (\pi/2)f_{PLL}$ (Hz). Jitter is spectral density $\times$ noise bandwidth, so

$$\langle\phi_e^2(V_n)\rangle = S_{vn}(\frac{K_V}{N\cdot f_{PLL}})^2 NBW_V = \frac{K_W}{N^2}\frac{\pi}{2f_{PLL}}\ (\text{rad}^2) \tag{3.18}$$

where K_W is the phase noise coefficient of the VCO.

Phase noise in the reference input is usually determined by the buffers that would follow a low-noise crystal oscillator. A known mean square jitter $\sigma^2_{\phi_{ref}}$ can be distributed across a constant spectral density up to $f_{ref}/2$ because whatever its actual (wide) bandwidth, it will be sampled by the phase detector at the rate of f_{ref}. This is transmitted to the input of the phase detector with a highpass function whose corner frequency is f_{PLL}, as in Fig.3.6(b). Usually $f_{PLL} \ll f_{ref}$, so the input of the phase detector sees almost the entire phase noise of the reference. Thus,

$$\langle\phi_e{}^2(ref)\rangle \simeq \sigma^2_{\phi_{ref}}\ (\text{rad}^2) \tag{3.19}$$

The loop returns quantization noise at the bang-bang phase detector's output to its input through a lowpass transfer function as in Fig.3.6(c). This follows by analysis on the graphs in this figure, or algebraically from (3.12) and (3.15) as

$$\begin{aligned}\frac{T}{1+T} &= \frac{1}{1+T^{-1}} \\ &= \frac{1}{1+\frac{s}{\omega_{PLL}}(1+\frac{s}{\omega_{P2}})\frac{1}{1+\frac{\omega_Z}{s}}} \\ &\simeq \frac{1}{(1+\frac{s}{\omega_{PLL}})(1+\frac{s}{\omega_{P2}})}\end{aligned} \tag{3.20}$$

The factorization in the last step assumes that the three roots are widely separated. This expression is of the lowpass form of (3.15a). From (3.15c), jitter due to quantization noise will thus be determined by a noise bandwidth of $NBW_Q = (\pi/2)f_{PLL}$.

Across the bandwidth NBW_Q which is $\ll f_{ref}$, quantization noise appears roughly with a uniform spectral density $(\pi-2)/\pi \times 2/f_{ref}$ from (3.7). Then starting from the second term on the right-hand side of (3.12), substituting the expression for quantization noise spectral density from (3.7), and multiplying by the noise bandwidth NBW_Q, we get

$$\langle \phi_e^2(q) \rangle = S_Q(f) \frac{1}{K_{BPD}^2} \cdot NBW_Q = (1 - \frac{2}{\pi}) \frac{\pi}{K_{BPD}^2} \frac{f_{PLL}}{f_{ref}} \text{ (rad}^2\text{)} \tag{3.21}$$

Using (3.18), (3.19), (3.21) and (3.5) the phase detector gain K_{BPD} may be found. The total phase noise at the input of the phase detector is the sum of the mean squares of three independent contributions given above. The VCO phase noise and the reference noise are uncorrelated because they arise from different sources. Quantization noise is uncorrelated to both, by definition as described in (3.3) and its surrounding text. Therefore, substituting (3.5) in the left-hand side of the following equation,

$$\sigma_{\phi_e}^2(total) = \sigma_{\phi_e}^2(VCO) + \sigma_{\phi_e}^2(\mathcal{Q}_{PD}) + \sigma_{\phi_e}^2(ref)$$

$$\text{we obtain } \frac{2}{\pi} \cdot \frac{1}{K_{BPD}^2} = \frac{\pi^2 K_W}{N\beta K_{BPD} V_{FS} K_V} + \frac{\pi - 2}{\pi} \cdot \frac{\beta V_{FS} K_V T_{ref}}{2N K_{BPD}} + \sigma_{\phi_{ref}}^2 \tag{3.22}$$

Solving this quadratic equation for K_{BPD}:

$$\boxed{\begin{aligned} K_{BPD} &= -\frac{\sigma_{\phi_0}}{2\sigma_{\phi_{ref}}^2} + \sqrt{\left(\frac{\sigma_{\phi_0}}{2\sigma_{\phi_{ref}}^2}\right)^2 + \frac{2}{\pi} \cdot \frac{1}{\sigma_{\phi_{ref}}^2}} \\ \text{where } \sigma_{\phi_0} &\triangleq \frac{\pi^2 K_W}{N V_{FS} K_V} \cdot \frac{1}{\beta} + \frac{\pi - 2}{\pi} \cdot \frac{V_{FS} K_V}{2N f_{ref}} \cdot \beta \end{aligned}} \tag{3.23}$$

The jitter parameter $\sigma_{\phi_0}^2$ may be interpreted with reference to (3.22) as the mean-square phase noise contributed by the VCO plus the quantization noise that appears at the input of the phase detector, if $K_{BPD} = 1$.

Since the effective phase detector gain K_{BPD} depends on the magnitude of the noise at its input, and since in turn this gain also determines f_{PLL}, bang-bang PLLs will display bandwidths that change with noise levels. There is a widely-held belief (e.g. [7]) that the effective gain K_{BPD} is always higher than the gain of a linear phase detector: we can now examine this quantitatively with the closed-form expression (3.23).

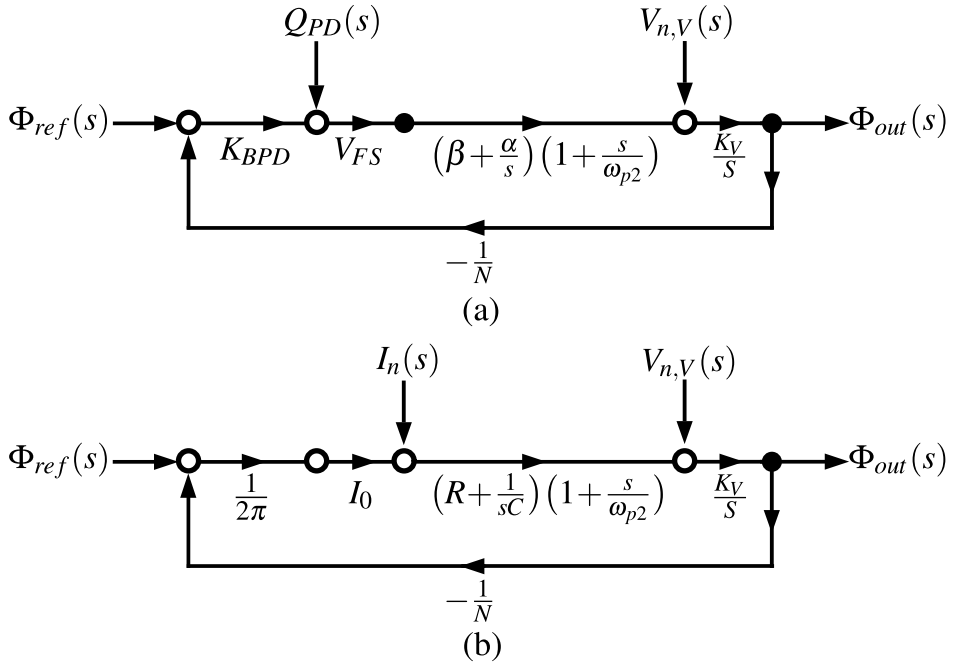


Figure 3.7: Comparing BB-PLL with CP-PLL; their respective signal flowraphs.

3.4 Bang-Bang PLL as Frequency Synthesizer

How does a bang-bang PLL compare with a charge-pump PLL if it is to realize the frequency synthesizer for a wireless receiver? While the synthesizer's output jitter affects the demodulated signal-to-noise ratio of the received waveform, the out-of-band phase noise spectral density is often more important since it limits the largest acceptable blocker. Thus, the design of a frequency synthesizer aims first to minimize the spectral density of out-of-band phase noise at certain critical offset frequencies while also holding jitter acceptably low. Without loss of generality, we compare phase noise roll-off and jitter of bang-bang and charge-pump PLLs, assuming for simplicity the same PLL bandwidth ω_{PLL} and the same K_W for the VCOs.

Quantization and the VCO typically contribute most of the noise at the output of a bang-bang PLL, whereas in a charge-pump PLL it is noise in the pump current and the VCO that are significant. As discussed in [92][93], the charge-pump noise will usually dominate the resistor noise in the loop filter. Fig. 3.7 shows the signal flowgraphs of the two loops under

consideration. The output jitter is found with a similar calculation as in Sec. 3.3.5:

$$\sigma^2_{\phi_{out,BB}} = \frac{\pi^2 K_W}{\omega_{PLL}} + S_{Q_{PD}}(f)\frac{N^2}{K^2_{BPD}}\frac{\omega_{PLL}}{4} \tag{3.23a}$$

$$\sigma^2_{\phi_{out,CP}} = \frac{\pi^2 K_W}{\omega_{PLL}} + S_{I_n}(f)\frac{N^2}{(I_0/2\pi)^2}\frac{\omega_{PLL}}{4} \tag{3.23b}$$

At high offset frequencies ($> f_{PLL}$), loop gain $|T(jf)|$ is $\ll 1$, thus

$$S_{\Phi_{out},BB}(f) = \frac{K_W}{f^2} + \left[S_{Q_{PD}}(f)\frac{N^2}{K^2_{BPD}}\frac{\omega_{PLL}}{4}\right]\frac{4\omega_{PLL}}{f^2} \tag{3.23c}$$

$$S_{\Phi_{out},CP}(f) = \frac{K_W}{f^2} + \left[S_{I_n}(f)\frac{N^2}{(I_0/2\pi)^2}\frac{\omega_{PLL}}{4}\right]\frac{4\omega_{PLL}}{f^2} \tag{3.23d}$$

Suppose that the two loops are designed to yield the same output jitter, as given by (3.23a) and (3.23b). Then (3.23c) and (3.23d) reveal that with the same bandwidth and power consumption in the VCO, a bang-bang PLL and a charge-pump PLL give the same phase noise. One loop does not have superior phase noise roll-off over the other when the output jitters are the same, and *vice-versa*. The trade-off between the quantization noise and VCO phase noise in a bang-bang loop becomes the same trade-off between charge-pump noise and VCO phase noise in a charge-pump loop. The output jitter, in both cases, contains a quadratic relationship in terms of ω_{PLL} (3.23a) and (3.23b), where there exists an optimum ω_{PLL} that minimizes the jitter. To lower the out-of-band phase noise, ω_{PLL} should be lowered to the point that the VCO phase noise becomes prominent.

In practical designs, K_W is limited by power consumption in the VCO, K_V is limited by complexity of oscillator tuning arrangements, V_{FS} is associated with the specific circuit and N is determined by the available reference frequency; design freedom remains only in the choice of loop filter gain β. Now we prove that this remaining design trade-off in bang-bang PLLs amounts to an optimization of β.

Using (3.23) and (3.6) and assuming the reference noise does not contribute significantly

to the output, jitter in a bang-bang PLL is

$$\begin{aligned}\sigma^2_{\phi_{out},BB} &= \frac{N^2}{K^2_{BPD}}\sigma_{\phi_0} \\ &= \frac{\pi}{4}\left[\sigma^2_{\phi_0} + \sigma_{\phi_0}\sqrt{\sigma^2_{\phi_0} + \frac{8}{\pi}\Big(N^2\sigma_{\phi_{ref}}\Big)^2}\right] \\ \text{where } \sigma_{\phi_0} &= \frac{\pi-2}{\pi}\cdot\frac{V_{FS}K_V}{2Nf_{ref}}\cdot\beta + \frac{\pi^2 K_W}{NV_{FS}K_V}\cdot\frac{1}{\beta}\end{aligned} \tag{3.24}$$

The output jitter scales monotonically with the parameter σ_{ϕ_0}, which itself depends quadratically on β. σ_{ϕ_0}, and thus the output jitter, are at a minimum when β_{opt} satisfies

$$\begin{aligned}\frac{\pi-2}{2\pi}\frac{K_V V_{FS}}{f_{ref}}\cdot\beta_{opt} &= \frac{\pi^2 K_W}{K_V V_{FS}}\cdot\frac{1}{\beta_{opt}} \\ \Rightarrow\ \beta_{opt} &= \sqrt{\frac{2\pi^3 K_W f_{ref}}{\pi-2}}\cdot\frac{1}{K_V V_{FS}}\end{aligned} \tag{3.25}$$

With (3.23) and (3.23c), the out-of-band phase noise spectral density of a bang-bang PLL can also be expressed in design parameters only:

$$S_{\Phi_{out},BB}(f) = \underbrace{\frac{K_W}{f^2}}_{\text{VCO}} + \underbrace{\frac{(\pi-2)K_V^2 V_{FS}^2}{2\pi^3 f_{ref}}\cdot\frac{1}{f^2}\cdot\beta^2}_{\text{Quant. Noise}} \tag{3.26}$$

As well as the VCO's phase noise which appears unaltered at the loop's output, there is quantization noise with a $1/f^2$ spectral density which rises with β^2.

The design process for a bang-bang PLLs is captured by (3.25) and (3.26), and may be summarized thus: for lowest out-of-band phase noise, one chooses the smallest $\beta \leq \beta_{opt}$ that still satisfies the jitter specification. This gives all told the best performance.

3.5 Adaptation in BB-PLLs

The expression (3.25) for optimum β was so far not known, nor that its value is determined by well-defined, repeatable parameters of a VCO circuit; these include the phase noise coefficient K_W, which is usually very repeatable chip-to-chip and changes weakly over operating temperature. Rather, it was believed that the optimum β is a noise-dependent parameter, and therefore in a practical design it must be found using some form of real-time adaptation.

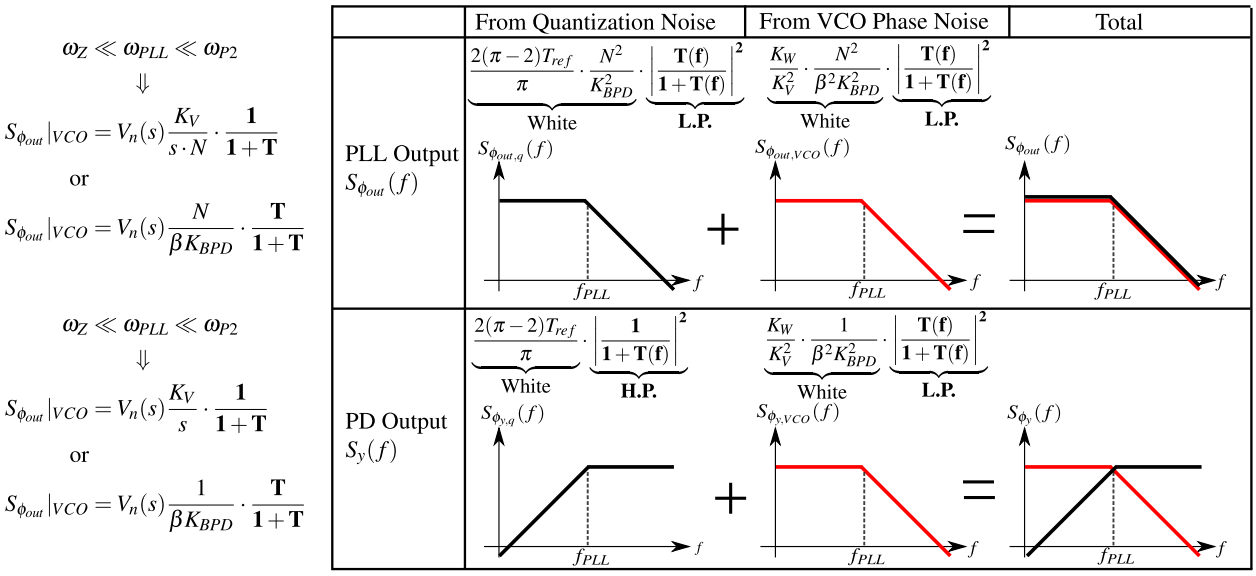


Figure 3.8: Contributions at optimum β to noise PSD $S_{\phi out}$ at PLL output, and to S_y at PD output, assuming $\omega_Z \ll \omega_{PLL} \ll \omega_{P2}$.

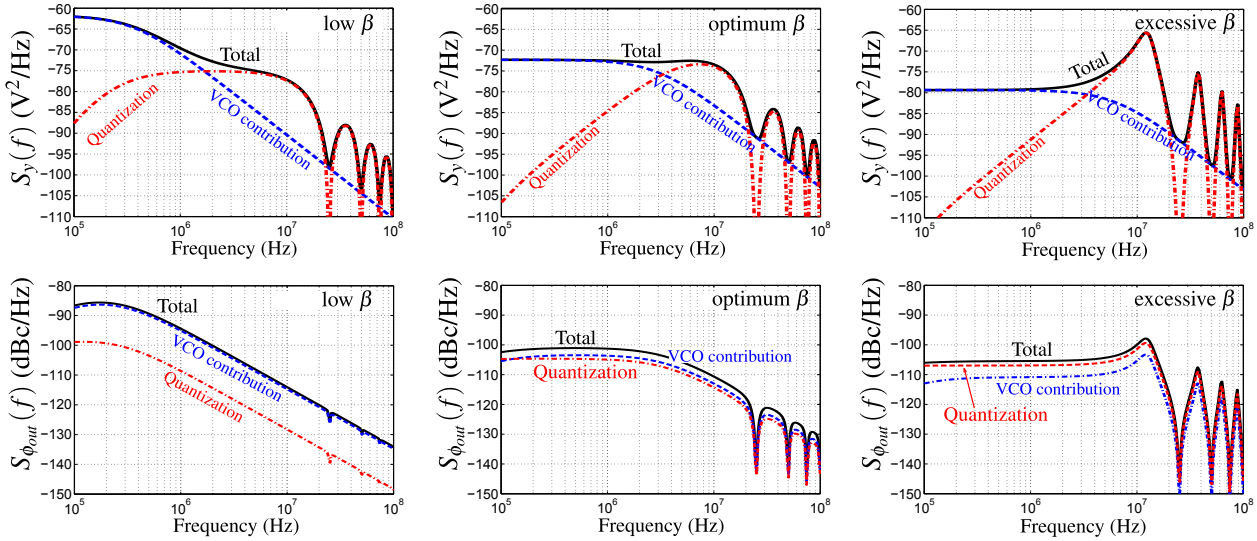


Figure 3.9: Example PSDs of S_y and $S_{\phi out}$ for a BB-PLL, showing relative contributions. (a) low $\beta = 0.13$; (b) optimum $\beta = 0.4$; (c) excessive $\beta = 1$. Parameters taken from [5], $T_{ref} = 40$ns, $N = 100$, $K_W = 300\text{dBc}\cdot\text{Hz}$, $\alpha = \beta \times 2^{-8}$, reference phase noise is ignored.

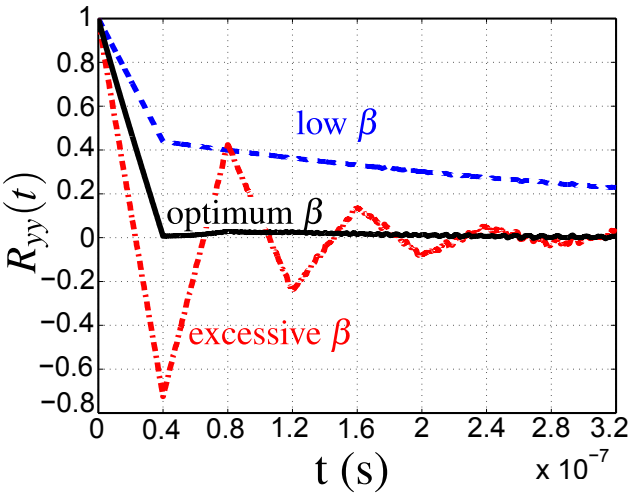


Figure 3.10: Autocorrelation $R_{yy}(t)$ of waveform at bang-bang phase detector output, when $\beta = 0.13,\ 0.4$ and 1 corresponding to the PSDs shown in Fig.3.9.

We have shown in (3.25) that in most cases the PLL can be designed to operate close to the optimum *without* adaptation.

Still, publications on adaptation raise some questions worthy of discussion that we can address here. In adaptation loops, jitter is monitored at the output $y(t)$ of the phase detector as a low frequency voltage (indeed it is what an instrument would display that measures phase noise on the VCO oscillation). For a BB-PLL $y(t)$ is a binary-valued square waveform clocked at f_{ref}. The injected quantization noise (Fig.7) propagates to the loop filter input through a highpass transfer function whose (lower) cutoff frequency is f_{PLL}. At β_{opt}, as Fig.3.8 shows, the contributions to the spectrum of $y(t)$ from the quantization noise and VCO phase noise will add to form a flat spectrum; this is verified by equating the expressions in the first row of Fig.3.8; the crystal reference in a frequency synthesizer usually contributes very little. Then from the second row it follows that $S_y(f)$ is constant up to $\approx f_{ref}/4$, or more precisely, that sample-and-hold action shapes it into a $\mathrm{sinc}^2(f/f_{ref})$ spectrum.

Departing from this optimum results in either VCO phase noise or the quantization noise becoming dominant at the PLL output. The PSD of $y(t)$ will no longer be constant across frequency if either VCO phase noise is emphasized at low frequencies or the quantization noise at high frequencies (Fig.3.8). At too large a β the corresponding autocorrelation function

$R_{yy}(t)$ will ring, while at a small β it will decay slowly over several reference cycles. These comparisons are shown as spectra in Fig.3.9 and as autocorrelation functions in Fig.3.10, where the effects of ω_Z and ω_{P2} are ignored because $\omega_Z \ll \omega_{PLL} \ll \omega_{P2}$ holds true over the various β. [77, 75] first pointed out that the output jitter is minimum when the contribution from quantization noise at the PLL output equals that from the VCO. Later [10, 5, 94] implemented real-time loop calibration to adjust β by measuring samples of $R_{yy}(t)$. But it seems that we are the first to furnish this simple proof of why jitter is at a minimum under these conditions.

As revealed by $R_{yy}(t)$, the spectral peak in quantization noise caused by excessive β might induce patterns in the '0' and '1's at the detectors's output. In this context [73, 75, 76, 74, 77, 5, 94] refer to a 'limit cycle regime'. It is true that in the absence of noise, the stable steady-state for any feedback loop with a binary element *must* be a limit cycle. The loop reaches a steady operating point when the binary element, which is metastable at zero input, settles into a toggling pattern whose average is the equilibrium magnitude (either digital or analog). However a large enough thermal noise will disrupt this limit cycle. The frequency spectra of waveforms in the PLL will no longer be discrete lines (as might be observed in noiseless simulations) but a continuous spectrum. [74] presents a simulation study of this phenomenon. The transfer functions we have investigated so far are able to explain observed continuous spectra, by modelling the fluctuations arising from the toggling phase detector as a quantization noise. We don't expect that limit cycles will appear in practical BB-PLLs operating under reasonable conditions, although it is easy to confuse highly peaked noise transfer functions in underdamped loops as evidence of a limit cycle. Our analysis offers a way not to fall into this trap. Section 3.7.3 discusses this with an actual example.

3.6 General Discrete Time Model to Study the Stability Limit

Loop stability can not be readily assessed if a PLL is modeled as purely a continuous time circuit. As first shown in [95] and later in [96], the second order analog PLL in Fig.3.11 is unstable at large loop gains. Since two poles can not (ideally) lead to instability, there

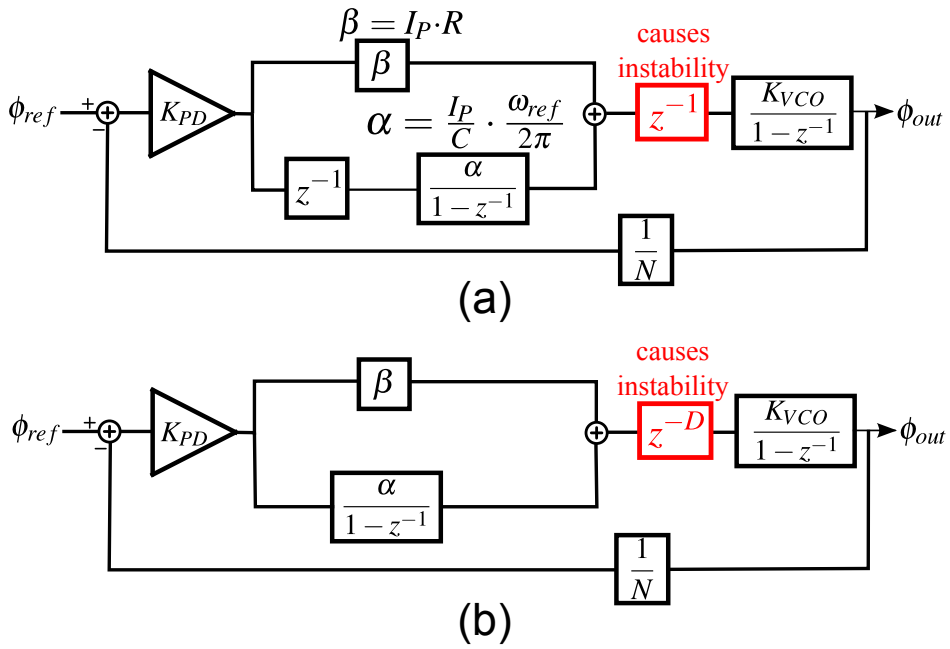


Figure 3.11: (a) Discrete time domain model for analog PLLs; (b) Discrete time domain model for digital PLLs.

must be an additional phase delay somewhere that has not been accounted for. This delay is inherent in the sampling action of any phase detector, which computes the phase error over a cycle of reference and then updates it. The loop delay inherent in the sampling operation, which is often overlooked in the continuous time model, poses an upper limit on loop gain for phase locked loops to remain stable. This is why although the continuous time model implies an unconditionally stable loop, delays added by sampling action and the use of digital filters will force the loop bandwidth to remain below some upper bound to ensure stability.

All frequency synthesis PLLs exchange information between two sampling frequencies, the (low) reference and the (high) oscillation frequency, modeled up-sampling and down-sampling operations. Exact discrete-time models to capture this may be complicated [75, 97], but to the first order, these operations may be approximated as averaging and multiplication as shown in Fig.3.11, in which operating frequency is normalized to the reference frequency. The focus of the proposed discrete time model is to discover the impact of key design parameters, at times with some sacrifices in accuracy. In analog PLLs, a loop delay of one reference

cycle is present because the VCO updates phase/frequency for the current cycle with the integrated voltage from the previous cycle. In digital PLLs, the loop filters may impose additional delays. With a total delay of D reference cycles, the general expression for loop gain in discrete time domain is then

$$T(z^{-1}) = K_{PD}\left(\beta + \frac{\alpha}{1-z^{-1}}\right)\frac{K'_{VCO}}{1-z^{-1}} \cdot \frac{1}{N} \cdot z^{-D} \tag{3.27}$$

where $z^{-1} = e^{-j2\pi f T_{ref}}$. The PLL will be unstable when $1 + T(z^{-1}) = 0$ has solutions $|z| > 1$. Oscillation in analog PLLs is called a *limit cycles* in digital PLLs. The highest stable bandwidth is set by the angle where the root locus crosses the unit circle in the z-plane. As $\beta \gg \alpha$, $\beta \gg \alpha/(1-z^{-1})$ at frequencies around $1 + T(z^{-1}) = 0$. $1 + T(z^{-1}) = 0$ is then simplified to

$$1 + \frac{\beta K_{PD} K_{VCO}}{N} \cdot \frac{z^{-D}}{1-z^{-1}} = 0 \tag{3.28}$$

The roots of $1 + T(z^{-1}) = 0$ must satisfy that $z^{-D}/(1-z^{-1})$ be real. When $D = 1$, this requires $z = -1$. When $D \geq 2$, it corresponds to $z^{-D} \approx j$. The stability limit is found by

$$\frac{\beta K_{PD} K'_{VCO}}{N} = \begin{cases} 2 & D = 1; \\ \dfrac{\pi}{2} \cdot \dfrac{1}{D} & D \geq 2. \end{cases} \tag{3.29}$$

K'_{VCO} is the normalized VCO gain in the discrete-time model and is related to K_{VCO} in the previous discussion as $K'_{VCO} = K_{VCO} \cdot T_{ref}$. Now we convert the stability limit to a more familiar form that

$$\frac{\beta K_{PD} K_{VCO}}{2\pi N} = \begin{cases} \dfrac{f_{ref}}{\pi} & D = 1; \\ \dfrac{1}{4} \cdot \dfrac{f_{ref}}{D} & D \geq 2. \end{cases} \tag{3.30}$$

For bang-bang PLLs, with (3.23) and (3.30), the stability limit in association with design parameters is given by

$$\frac{\beta K_{VCO}}{2\pi N}\left(-\frac{\sigma^2_{\phi_0}}{2\sigma^2_{\phi_{ref}}} + \sqrt{\left(\frac{\sigma^2_{\phi_0}}{2\sigma^2_{\phi_{ref}}}\right)^2 + \left(\frac{2}{\pi}\right)^2 \cdot \frac{1}{\sigma^2_{\phi_{ref}}}}\right) < \begin{cases} \dfrac{f_{ref}}{\pi} & D = 1; \\ \dfrac{1}{4} \cdot \dfrac{f_{ref}}{D} & D \geq 2. \end{cases} \tag{3.31}$$

The stability limit is better visualized by the root locus plots in Fig.3.12, where $K_P = \beta K_{PD} K_{VCO}$ is swept from zero to infinity in each plot. The highest achievable bandwidth is set by the angle where the roots cross the unit circle. With zero loop delay $D = 0$, the PLL is unconditionally stable. The two poles never cross the unit circle. The root locus plot in Z-domain under zero loop delay in Fig.3.12(a) provides the same conclusion as the continuous-time model in S-domain. However, zero loop delay is a fiction. The PLL becomes conditionally stable when the loop delay D is increased from 0 to 1. When K_P is low, the roots locate closely to $z = 1$, leading to narrow bandwidths and peaking in the output phase noise spectrum. With increasing K_P, the roots diverge away from the $z = 1$. The longer distance between the roots and the unit circle suppresses the peaking and widens the bandwidth. Finally, when K_P becomes too large, one of the roots would approach the unit circle at $z = -1$ and finally exceed the unit circle, leading to instability. Strong peaking in spectrum will be observed again when the loop approaches or exceeds its stability limit. When the loop delay D is further increased to 2 or 3, poles escape the unit circle at lower gains. The achievable bandwidth diminishes severely with increasing loop delay. Peaking will appear in the output phase noise spectrum when K_P is either too low or too high, but for very different reasons. The optimal K_P, and thus the PLL bandwidth, is defined between the two extremes. If a wide loop bandwidth is being sought, digital PLLs have the disadvantage of being vulnerable to *limit cycles*. Digital PLLs are of interest for wideband applications because it is believed that digital circuits keep abreast of technology scaling. But the more stringent requirements on stability are at odds with high loop bandwidths.

Compared with previous work [73, 75, 66], (3.31) offers a more straightforward way to check whether a design satisfies the stability limit requirement. It also clarifies the long standing debate on *limit cycles*, an undesired phenomenon when the quantization noise dominates over other thermal noise sources as discussed in [71, 11, 7, 73, 75]. According to these works, the bang-bang phase detector can be linearized when the input jitter is below some threshold. This mysterious range is referred as the 'random noise regime'. The effective gain is then calculated by the probability density function (PDF) of the input jitter. When the input jitter is larger than the threshold, mainly due to the large quantization noise, the

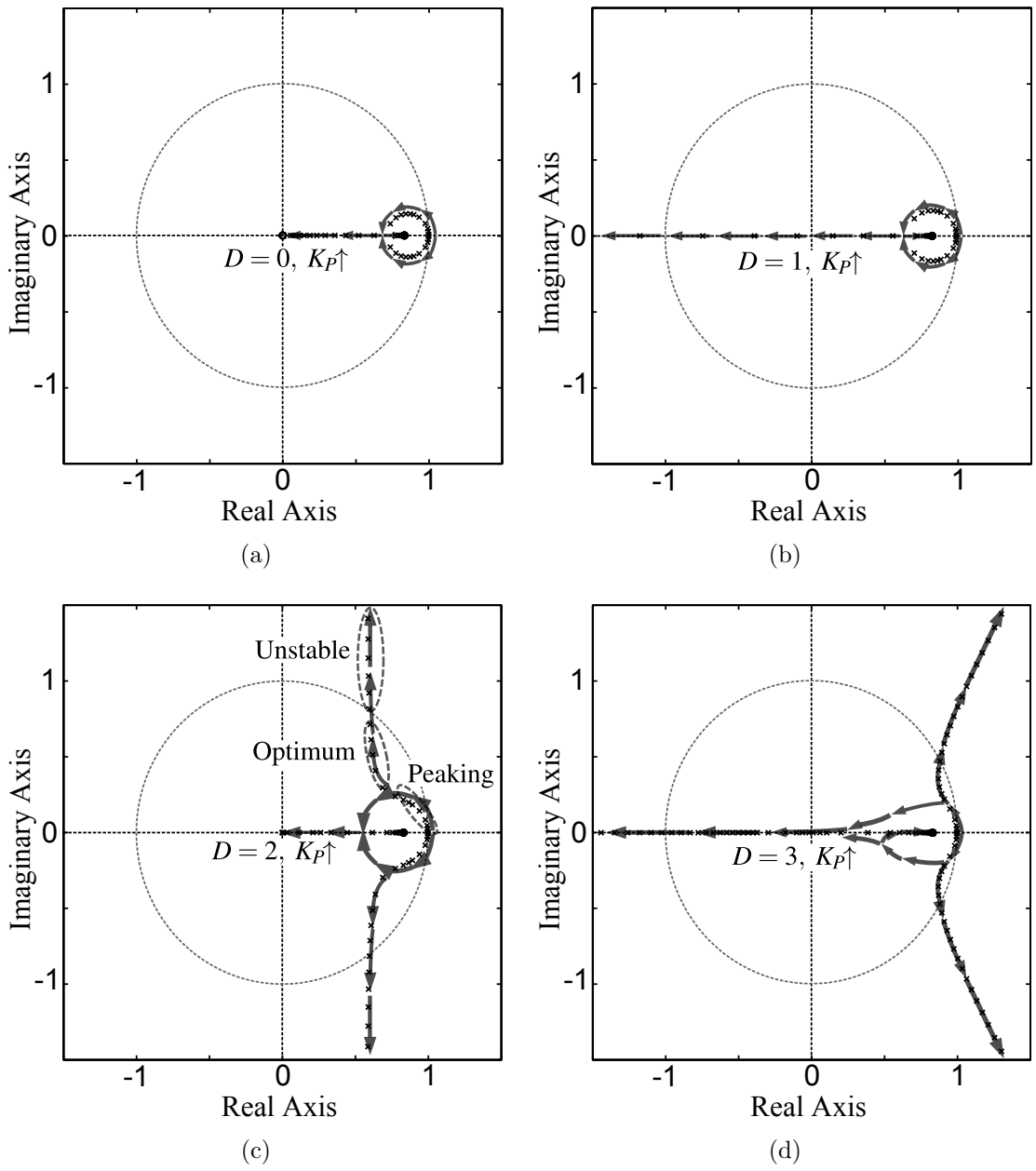


Figure 3.12: Root locus plots with the discrete-time model in Fig.3.11 under various loop delays, $K_P = \beta K_{PD} K_{VCO}$, (a) $D = 0$; (b) $D = 1$; (c) $D = 2$; (d) $D = 3$. Simulation parameters: $T_{ref} = 15$ns, $N = 40$, $\beta/\alpha = 5$.

bang-bang phase detector can no longer be linearized, resulting in limit cycles. This is also referred as 'frequency granularity' in [95]. A common remedy to avoid limit cycles is to dither the reference as in [9, 98, 55]. Our analysis incorporates these observations into a unified

frame. With excessive β or K_{VCO}, quantization noise will dominate K_{BPD} in (3.23) and it becomes likely that the stability limit in (3.31) will be violated. With large noise, the effective gain K_{BPD} is low. Thus, limit cycles caused by instability is usually accompanied by a high in-band phase noise. Raising noise of the reference $\sigma_{\phi_{ref}}$ further degenerates the effective gain of the bang-bang phase detector, which eliminates limit cycles but at the cost of the extra phase noise at the output. Since the transfer function from the reference to the input of phase detector is high-pass and to the output of the PLL is low-pass, $\Sigma - \Delta$ modulation the reference phase noise spectrum achieves the same goal without introducing extra phase noise. It shifts low-frequency phase noise to high frequencies. Thus, most of the reference noise appears at the input of the bang-bang phase detector to degenerate the detector gain while PLL output phase noise is not affected because of the low-pass transfer characteristic to the output. This has been explored in TDC-based PLLs [98, 55, 9]. However, (3.23) and (3.31)) suggest a simpler solution. Reducing β increases the phase noise contribution from the VCO at the input of bang-bang phase detector, thus degenerates the effective gain K_{BPD}. Reducing β also alleviates the requirement in (3.31). So stable operation can be retained by reducing β. The idea of reducing β is already implemented in [11, 7] and referred as 'intrinsic noise dithering by using VCO thermal noise'. We have now precisely explained what this means.

3.7 Verification against Measurements

We have validated the analysis developed so far against published measurements. Using small sets of reported design parameters with the newly derived phase detector gain in (3.23), we are able to calculate loop transfer functions. This enables rapid calculation of the output spectrum without resort to complicated simulation. Further we are able to identify the relative contributions to phase noise and jitter, and the effects of changing loop parameters.

3.7.1 Case 1: Sensitivity to Reference Noise

[6] reports measurements on a bang-bang PLL of the output phase noise spectrum as reference phase noise rises; this is shown in Fig.3.13. This work was modeling clock recovery from inputs corrupted by increasingly larger amounts of jitter. A signal generator synthesizes a reference periodic waveform with programmable amounts of jitter. At low levels of reference jitter, the in-band phase noise at the PLL output does not track rises in reference noise, the bandwidth of the loop remains almost unchanged, and the spectrum is relatively flat. Once the reference jitter exceeds a certain level, however, the in-band phase noise at the PLL output starts to track the reference noise and the bandwidth of the loop shrinks; and a peak in the jitter spectrum becomes increasingly prominent. Although [6] does not explain why, the analysis in Section 3.3 does.

Fig.3.13 plots spectral density as predicted by our expressions. It matches measurement results remarkably well across the full sweep of added jitter.

3.7.2 Case 2: Changing Parameters of Loop Filter

[7] shows that by tuning the loop filter parameters β and α, without affecting in-band phase noise, the bandwidth of a bang-bang PLL may be adjusted. Our analysis shows that when quantization noise dominates within the loop passband, the in-band phase noise will remain constant as β and α change. A fixed ratio between β and α will fix the zero frequency ($\omega_z = (\alpha/\beta) f_{ref}$), ensuring enough phase margin as the loop gain drops to prevent peaking.

Using reported design parameters and (3.23), we calculate that the parameter σ_{ϕ_0} in [7] ranges between 1.5×10^{-5} and 6.9×10^{-5} rad, which is $\ll \sigma_{\phi_{ref}}$ of a typical 40MHz crystal oscillator used for the loop's reference. From (3.23) we may conclude that now the phase detector gain K_{BPD} will be determined by the reference noise only (the third term). Using (3.16) and (3.23), $\sigma_{\phi_{ref}}$ is extracted from measurement Fig.3.14(a). Spectra calculated with this extracted $\sigma_{\phi_{ref}}$ are plotted in Fig.3.14(b) and Fig.3.14(c). They match the measured total spectrum well across a range of β, α, except for flicker noise at low frequency offsets.

[7] also reports that the phase noise for $\beta = 0.023$ is -139dBc/Hz at 20MHz frequency

T_{ref}	15ns
N	20
K_V	$2\pi \times 10^6$rad/sec-V
β	0.2
βK_V	$0.2 \times 2\pi \times 10^6$rad/sec-V
α	$\beta/10$
K_W	5×10^3rad$^2 \cdot$Hz
V_{FS}	1V
$\sigma_{\phi_{ref}}$ (Typical)	$1 \times 10^{-4} \sim 4 \times 10^{-2}$rad

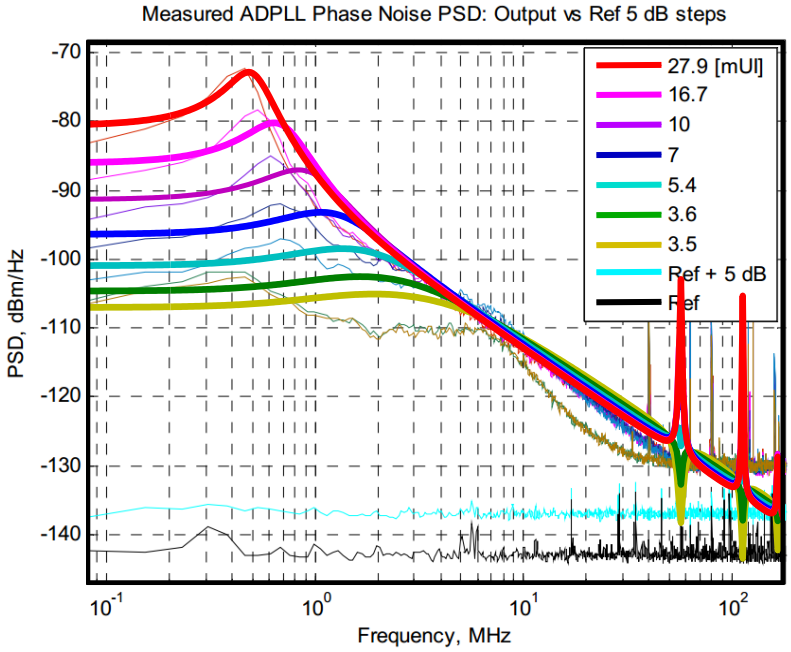
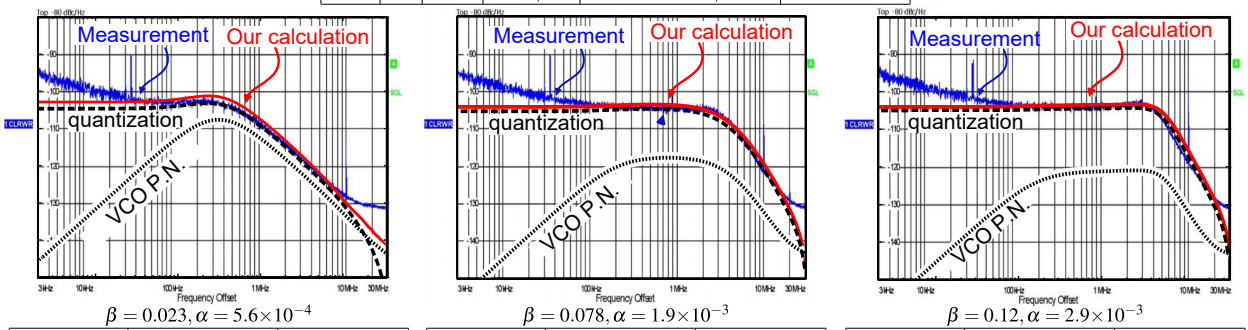


Figure 3.13: Measured output phase noise spectrum (thin lines) from [6, Fig. 4] for rising levels of reference jitter, compared with calculated output phase noise spectrum (thick lines).

offset. Although [7] does not say so, we will assume that this was found to be the optimum value of this parameter. Our analysis proves that it is. At this value of β and α, quantization noise and VCO phase noise contribute *equal* amounts at 20MHz offset. When β is increased to 0.078 and further to 0.12, the phase noise at 20MHz frequency offset worsens to -130dBc/Hz and -127dBc/Hz respectively (neither value reported in [7]). According to our analysis, for these parameter values, the quantization noise will dominate the VCO's phase noise at 20 MHz offset (see Fig.3.14), resulting in a sub-optimal outcome.

The usefulness of the analysis presented in this paper is now clear for this frequency

T_{ref}	N	V_{FS}	K_W	K_V	$\sigma_{\phi_{ref}}$(Typical)
25ns	90	0.6V	5rad^2/Hz	$2\pi \times 3 \times 10^6$rad/sec-V	7×10^{-4}rad



$\beta = 0.023, \alpha = 5.6\times10^{-4}$

	Our Calculation	Measurement
Bandwidth	530KHz	500KHz
In-band phase noise	-102dBc/Hz	-102dBc/Hz
Integrated Jitter	390ps	460ps

(a)

$\beta = 0.078, \alpha = 1.9\times10^{-3}$

	Our Calculation	Measurement
Bandwidth	3.1MHz	4MHz
In-band phase noise	-103dBc/Hz	-103dBc/Hz
Integrated Jitter	660ps	793ps

(b)

$\beta = 0.12, \alpha = 2.9\times10^{-3}$

	Our Calculation	Measurement
Bandwidth	4.6MHz	5.5MHz
In-band phase noise	-103dBc/Hz	-103dBc/Hz
Integrated Jitter	880ps	973ps

(c)

Figure 3.14: Measured output phase noise spectra taken from [7, Fig. 14] with swept loop parameters, compared with calculated spectra. $T_{ref} = 25$ns, $N = 90$, $K_{VCO} = 3$MHz/V, $\beta = 0.023, 0.078, 0.12$, $\alpha = [5.6\times10^{-4}, 1.9\times10^{-3}, 2.9\times10^{-3}]/T_{ref}$, $\phi_{ref} = 50\mu$UI, VCO phase noise index $K_W = 5\text{rad}^2 \cdot \text{Hz}$ obtained from [8] with $I = 1.5$mA, $V_{DD} = 1.2$V, $f_0 = 3.6$GHz and inductor $Q = 10$.

synthesizer loop: it straightforwardly reveals the relative spectral densities responsible for the total output phase noise and their dependence on loop design parameters, and it guides the designer towards making the optimal choice, thereby minimizing design by trial-and-error and lessening reliance on simulations.

The quantization noise of the fractional-N divider in this PLL is cancelled by a specially designed circuit [7]. In Section 3.9 of this paper we will examine the consequences if this circuit is removed.

3.7.3 Case 3: Investigating Stability

Using behavioral simulations, [9] probes the stability limit of the bang-bang all-digital phase locked loop (Fig. 3.15). When β and α are both increased by 20× in a digital loop filter with zero latency, a higher bandwidth is obtained with no peaking. However, when a loop

T_{ref}	3.6ns
N	40
$K_P(\beta K_V)$	$2\pi \times [0.5, 10] \times 10^6 \text{rad/V} \cdot \text{s}$
K_W	$1.3 \times 10^3 \text{rad}^2 \cdot \text{Hz}$
V_{FS}	0.5V
$\sigma_{\phi_{ref}}$ (Typical)	3×10^{-3}rad

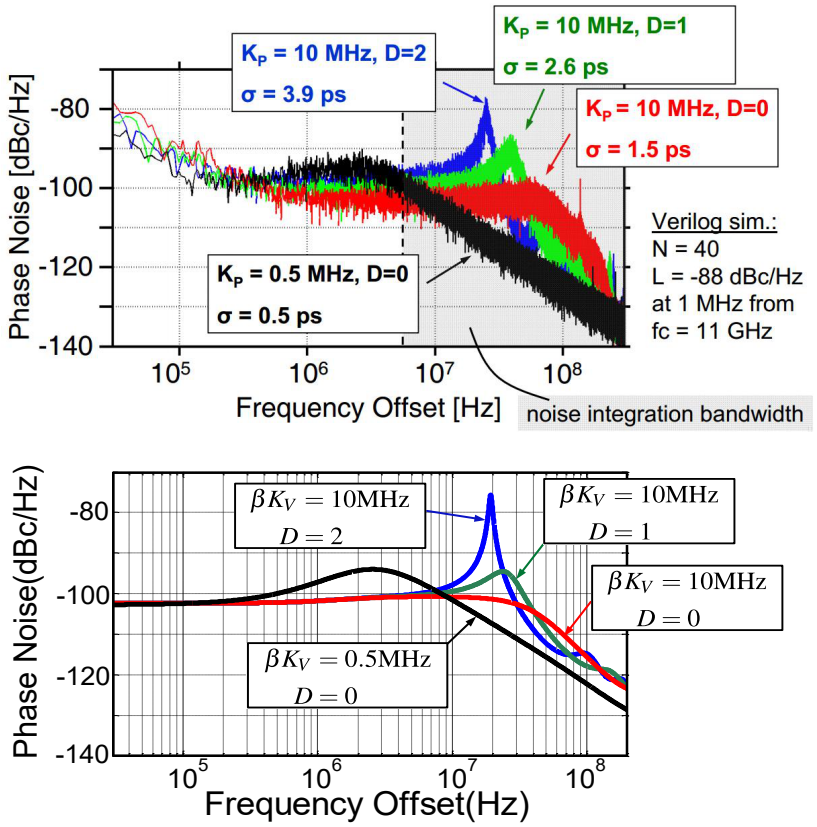
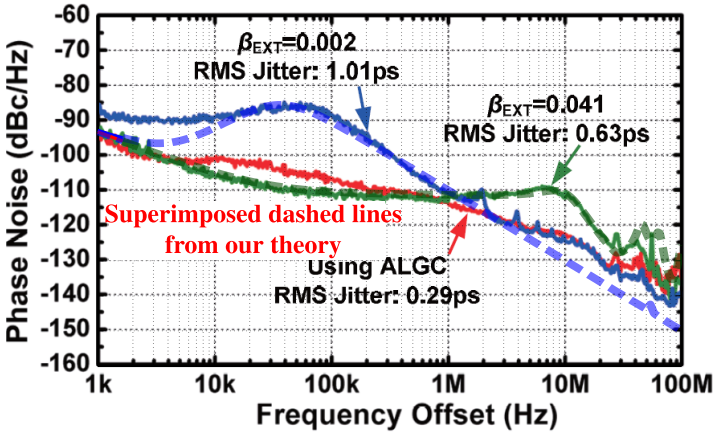


Figure 3.15: Output phase noise spectrum with various loop gain and loop delay from [9, presentation slide 5] versus our calculations.

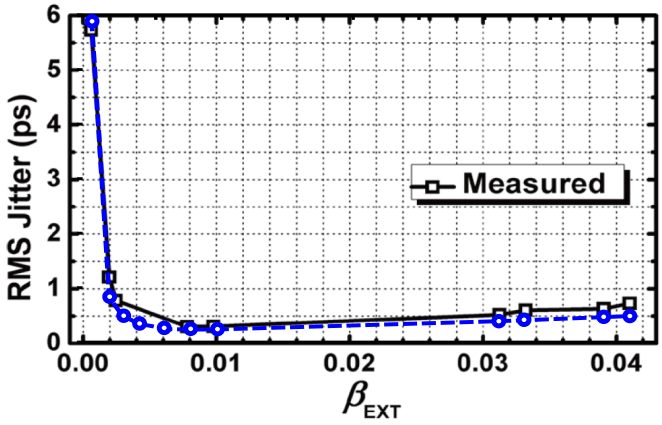
delay, first, of one clock cycle ($D = 1$) is added, and then of two clock cycles, the loop tends towards instability showing a prominent peak in the output phase noise.

In this case, we will also assume that $\sigma_{\phi_{ref}}$ dominates σ_{ϕ_0} at the phase detector's input. $\sigma_{\phi_{ref}} = 3 \times 10^{-3}$ rad is extracted from the simulated phase noise spectrum when $K_P = 10\text{MHz}$ and $D = 0$. The jitter parameter σ_{ϕ_0} changes minimally over the interval $[0.52 \times 10^{-3}, 0.54 \times 10^{-3}]$ rad with changing β, but remains $\ll \sigma_{\phi_{PLL}}$. Although reference noise dominates at the input of the phase detector, it will be greatly attenuated at the loop output by the low-pass

T_{ref}	18.2ns
N	72
α	$\beta/512$
K_V (Typical)	$2\pi \times 3 \times 10^6$rad/sec-V
K_W	$10\text{rad}^2 \cdot \text{Hz}$
V_{FS}	1V
$\sigma_{\phi_{ref}}$ (Typical)	9×10^{-5}rad



(a)



(b)

Figure 3.16: (a) Measured output phase noise spectrum (solid) with various β from [10, Fig. 4] versus calculated output phase noise spectrum (dashed). Dashed lines may be hard to discern because they overlay measured curves so accurately; (b) Measured output jitter (solid with square markers) for various β [10, Fig. 4] versus our calculated output jitter (dashed with round markers).

transmission function. Our analysis predicts the simulation-based results of [9] accurately (Fig.3.15). To account for loop delays, we employ a discrete-time PLL model as in [95].

That we can use transfer function analysis to predict the sharp (20 dB) peak when $D = 2$ proves that this peak is not due to a limit cycle. It is due to noise alone, and must not be confused for a discrete spectral line. In this way, our analysis serves as a useful tool to rule out suspected limit cycles during loop simulation and measurement.

3.7.4 Case 4: Loop Calibration

[94, 10] describe similar schemes to adapt loop bandwidth, that automatically adjust β based on the autocorrelation of phase detector's output, as described in Section V.

Fig.3.16(a) shows that in-band phase noise is higher when $\beta = 0.002$ compared to when $\beta = 0.041$, so we conclude that quantization noise dominates the in-band phase noise. This in-band phase noise depends only on the divider ratio N and K_{BPD}. Using (3.13), we first calculate the effective gain K_{BPD} when $\beta = 0.002$. K_V is then extracted using (3.16) with a loop bandwidth of 200KHz as shown in Fig.3.16(a). $\sigma_{\phi_{ref}}$ is extracted in the same way as in Cases 2 and 3, by assuming K_{BPD} is dominated by $\sigma_{\phi_{ref}}$ when $\beta = 0.041$. With $K_V = 3$MHz/sec-V and $\sigma_{\phi_{ref}} = 9 \times 10^{-5}$rad, our calculation matches the measurement in Fig.3.16(a). These assumptions are verified further by comparing $\sigma_{\phi_{ref}}$ and σ_{ϕ_0}, when $\beta = 0.002$: $\sigma_{\phi 0} = 1.2 \times 10^{-4}\text{rad} > \sigma_{\phi_{ref}}$, implying the VCO phase noise dominates over quantization noise to lower the loop gain. But when $\beta = 0.041$, $\sigma_{\phi 0} = 2.8 \times 10^{-5}\text{rad} < \sigma_{\phi_{ref}}$ since the larger β now reduces the contribution from the VCO phase noise.

Fig.3.16(b) plots our predictions of integrated jitter versus varying β against the measurement. Using extracted K_V and $\sigma_{\phi_{ref}}$, our calculation matches measurement accurately. The optimum β is found when the VCO phase noise balances the reference and quantization noise at the PLL output. Departing from this optimum results in either a dominant VCO phase noise when β is lower, or quantization noise at high β.

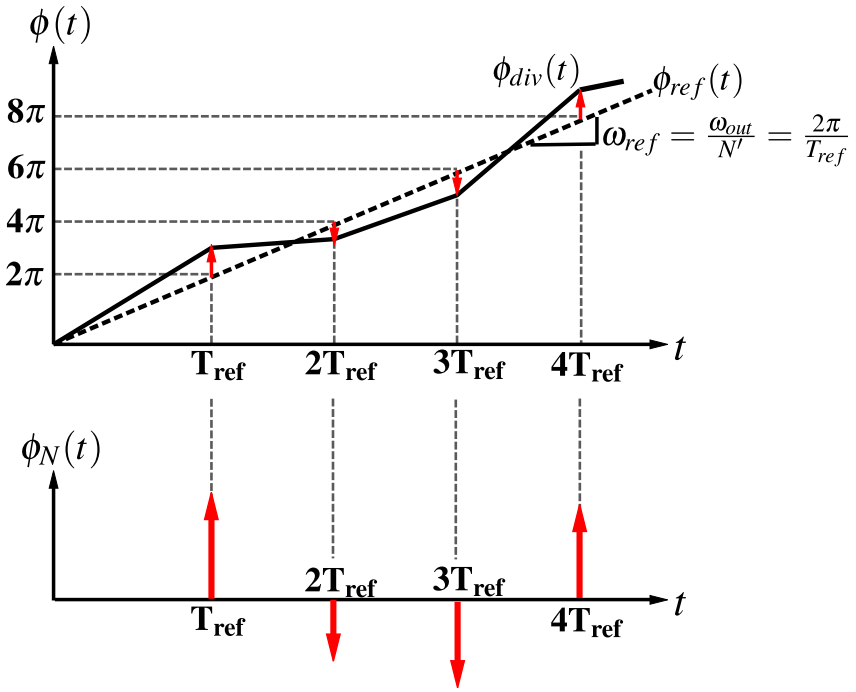


Figure 3.17: Phase trajectory and fractional divider-induced phase error sampled every T_{ref}.

3.8 Loop Response to Fractional-N Divider

Assume that the loop is synthesizing $f_{out} = N' f_{ref}$, where N' is non-integer. This divide ratio is the average value of the divider's modulus, as it is toggled between integers by a delta-sigma modulator of order m. After passing through the lowpass transfer function to the loop's output, most of this quantization noise will be filtered out.

Fractional-N quantization noise poses a unique problem in a BB-PLL, because it is transmitted to the phase detector's input through a highpass transfer function, and therefore appears there almost in its entirety. This easily dominates the thermal noise sources as well as the phase detector's own quantization noise, thereby lowering the phase detector gain K_{BPD}. With the loop gain thus weakened, the output phase noise will rise. We can now predict just by how much.

The quantization noise from the divider $\Phi_N(s)$ injects into the loop's signal flowgraph (Fig. 3.4) effectively at the same node as the reference phase noise $\Phi_{ref}(s)$. Every toggle of the divider modulus abruptly changes the instantaneous frequency, given by the slope of the

T_{ref}	25ns
N	90
K_V	$2\pi \times 3 \times 10^6 \text{rad/V} \cdot \text{s}$
K_W	$3\text{rad}^2 \cdot \text{Hz}$
β, α	$1.4 \times 10^{-3}, 2.4 \times 10^{-5}$
V_{FS}	0.6V
$\sigma_{\phi_{ref}}$	0

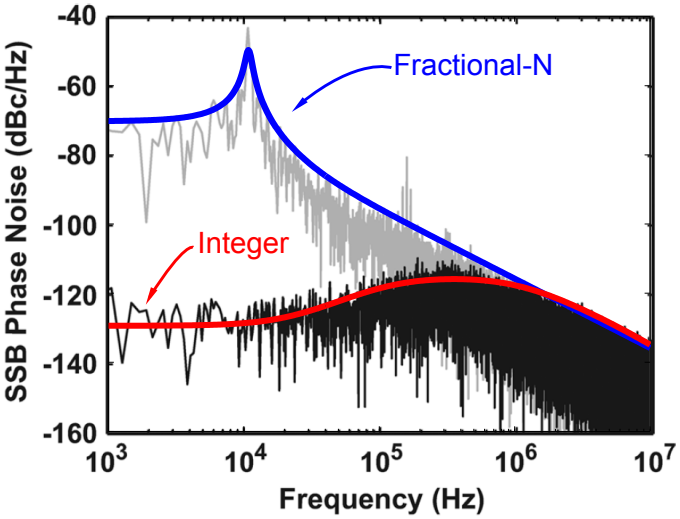


Figure 3.18: Calculated output phase spectra versus simulated spectra in [11], assuming first-order Σ-Δ modulation.

divided output's phase trajectory[2]. Since the sigma-delta modulator scrambles the toggling sequence, its action may be modelled in the phase domain by random jumps in phase at the input of a (hypothetical) divider with constant non-integer modulus N', as shown in Fig.3.17. The resulting mean-square phase error at the output of fractional-N divider depends on the order of Σ-Δ modulation. As derived in the Appendix,

$$\langle \phi_N^2 \rangle = \begin{cases} \left(\frac{2\pi}{N'}\right)^2 \times \frac{1}{12} & \text{for} \quad m = 1 \\ \left(\frac{2\pi}{N'}\right)^2 \times \frac{1}{12} \times 2 & \text{for} \quad m = 2 \\ \left(\frac{2\pi}{N'}\right)^2 \times \frac{1}{12} \times 6 & \text{for} \quad m = 3 \end{cases} \tag{3.32}$$

It is rare that $m > 3$ in practice.

[2] [90, Fig. 5] defines and illustrates the phase trajectory of an oscillator.

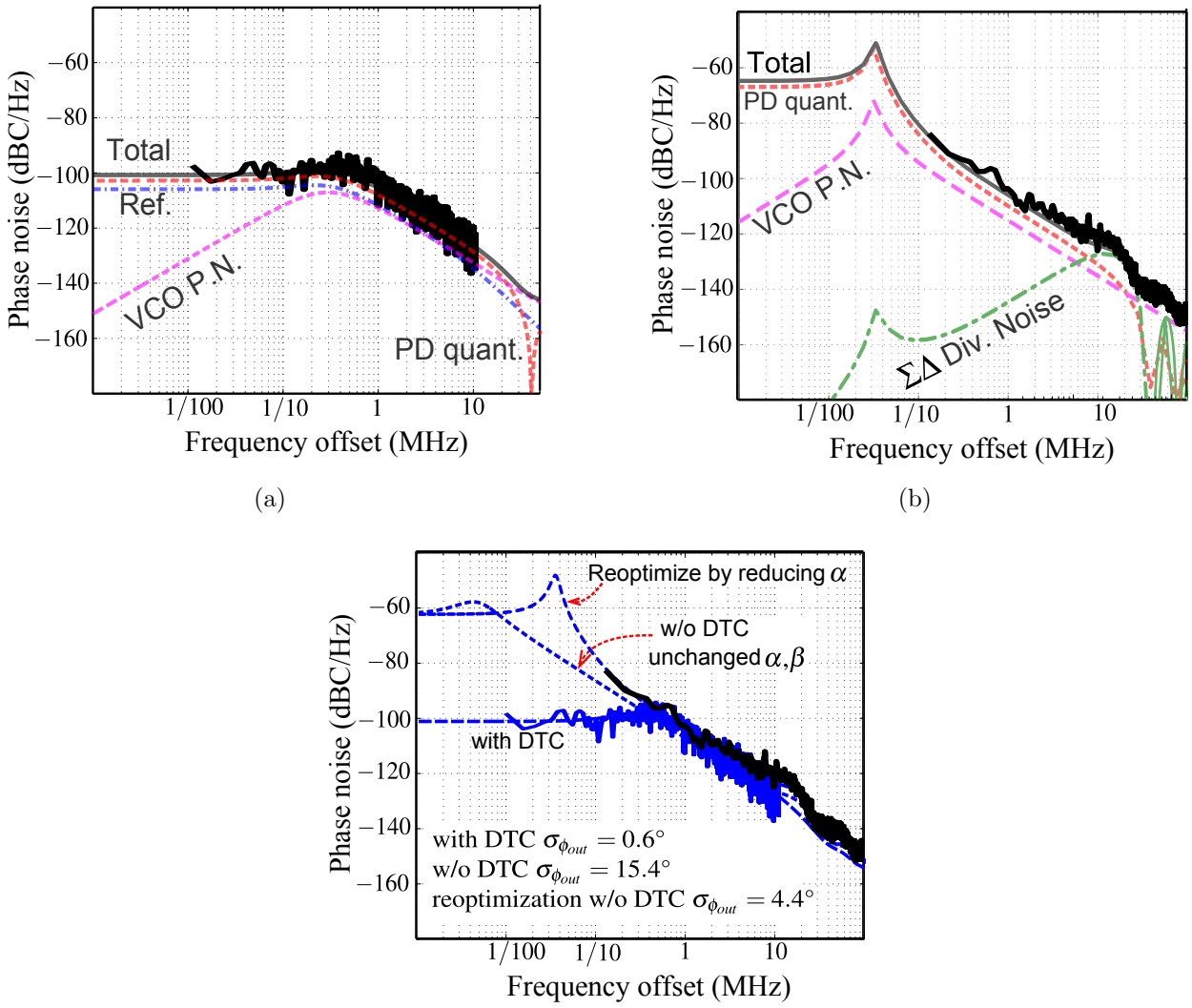


Figure 3.19: Effects of including shaped quantization noise from frac-N divider in PLL of [7]. $\beta = 0.023$, $\alpha = 5.6 \times 10^{-4}$ and other parameters from Fig.3.14. Predicted PLL output phase noise versus event driven simulation (solid black lines) when (a) divider noise from 3^{rd} order $\Sigma\Delta$ divider noise is cancelled (absent); (b) when it is present; (c) after β, α are re-optimized for minimum jitter.

The Σ-Δ modulator shapes the spectral density of this quantization noise to a null at DC and a maximum at the Nyquist frequency of $\frac{1}{2}f_{ref}$. From (3.12), this is transmitted to the phase detector input through a highpass transfer function. The lower corner frequency f_{PLL} is $\ll f_{ref}$, which means that $\langle\phi_N^2\rangle$ appears almost in its entirety at the phase detector input. It is simple to show that this is a great deal larger than the jitter appearing there from all other sources of injected noise. So the phase detector gain is now

$$K_{BPD} \simeq \sqrt{\frac{2}{\pi}} \frac{1}{\sigma_{\phi N}} \tag{3.33}$$

Frac-N noise lowers the phase detector gain without appearing noticeably in the spectrum at the loop output, because its spectral density has been shaped to remain lower than all other contributions across the noise bandwidth of $\sim f_{PLL}$. This stealthy effect must be modelled properly.

We have shown that the optimum choice of the key loop filter coefficient, β, minimizes output mean-square jitter. The optimum β is independent of divider modulus and remains unchanged in the presence of a frac-N divider.

In our earlier discussion of the loop in [7], we had neglected the frac-N noise. This is because that frequency synthesizer includes a digital-to-time converter (DTC) driven by the output of the delta-sigma modulator to reconstruct the discontinuities in the phase trajectory at modulus transitions. In effect, the DTC subtracts the frac-N noise and the loop perceives a time-invariant divider of non-integer modulus N'.

But like any other circuit at the digital-analog boundary—especially when that analog variable is time—the design of the DTC is difficult. So the question arises: how would the loop perform if the DTC were not used? Our theory provides a quick answer, as illustrated in Fig.3.18 and Fig. 3.19. The authors reporting the spectrum in Fig.3.18 have used time-domain simulations to compare the PLL output spectrum of an integer-N PLL (black solid lines) and the spectrum of a fractional-N PLL without cancelling divider noise (grey solid lines). Using parameters from [11] and transfer functions, we are able to predict the output spectrums changes accurately. In Fig. 3.19, the spectrum labelled "w/o DTC" appears when the full frac-N quantization noise circulates in the loop, and it may be compared with

the spectrum labelled "with DTC" when the frac-N noise has been subtracted. All loop parameters are held the same. Also plotted on these two curves are the results of event-driven simulations of both loops in SIMULINK, which establish that the analytical expressions we have used are accurate. Whereas the results of the analysis can be plotted in minutes, the simulation consumes 8 to 10 hours of computing time, and even then it can span only a fraction of the upper frequency range.

There are two features in this comparison worthy of note: 1) In the frequency range where both spectra have a slope of $-20\,\mathrm{dB/decade}$, the spectral densities are essentially equal, except for a small bump between 10 and 20 MHz; 2) Frac-N noise raises the flat region of the spectrum by 30 dB. This is entirely because of the lowered K_{BPD}, which reduces loop gain. Since the coefficients α and β of the loop filter are unchanged, the smaller K_{BPD} can pull the unity frequency of the loop gain *below* the zero frequency, drastically shrinking phase margin. Calculations using (3.8) and (3.9) show that this is indeed the case here, and that the phase margin is about 15°, which predicts a spectral peak of roughly $1/(2\sin(\frac{1}{2}15°)) = 9\,\mathrm{dB}$, close to the simulated peak.

What, then, are the consequences of removing the DTC? If the only criterion were blocker tolerance at 20 MHz offset, then there is a small penalty exacted by the spectral bump. However if mean square phase noise, or jitter, matters, as it does increasingly when demodulating multi-point signal constellations, then removing the DTC can raise the jitter substantially. A reasonable way to quantify this is after adjusting the loop filter coefficients so as to minimize jitter. Fig. 3.19(c) shows the result of this re-optimization, where the peak in output spectrum is no longer prominent, but still the jitter remains almost 10× larger than it is with the DTC present. Does this mean that a frac-N frequency synthesizer with a BB-PD is unusable for any realistic use, unless the quantization noise is subtracted? By comparing [99, Fig. 18] with Fig. 3.19(c) we see that the phase noise without DTC meets Bluetooth specifications. Whereas [99] describes a frac-N loop that employs a 24-tap time-to-digital converter (TDC) to detect phase [100], we have shown that a loop with a binary phase detector can satisfy the needs of a Bluetooth receiver. The design of the BB-PLL is undeniably a great deal simpler. Its higher RMS jitter of 4.4° remains acceptable for GSM[101].

In a charge-pump fractional-N PLL [101], high order Σ-Δ modulation is used to suppress the in-band divider quantization noise within the loop bandwidth. This is not necessarily needed in a bang-bang fractional-N PLL, where in-band phase noise is dominated by the quantization noise from the phase detector. (3.32) and (3.33) reveal that the detector gain K_{BPD} degrades with the order of Σ-Δ modulator increasing, as described qualitatively in [11]. Thus, minimizing jitter in a bang-bang PLL requires choosing Σ-Δ modulation of suitable order to suppress in-band frac-N noise without over compromising the detector gain K_{BPD}. The integrated jitter of 4.4° in Fig.3.19 was calculated assuming a 3rd order Σ-Δ modulator driving the divider, but, with a 2nd order Σ-Δ modulator, jitter drops to 3.1°. This is the optimal modulator order. Reducing the order to one would raise output jitter to 3.6°, when the frac-N noise begins to raise in-band phase noise.

What are the options if a frac-N BB-PLL synthesizer is to be designed for low jitter? Our analysis shows that the frac-N noise from a modulator of suitable order does not contribute much to the final jitter itself; instead it does damage by lowering the phase detector gain, thereby loop gain, so that other sources of noise—in this example, the quantization noise of the BBPD—raise the output jitter. There are three remedies to this: 1) Subtract the frac-N noise with a DTC. It is difficult to design this circuit block; one solution uses LMS adaptation to obtain the correct time delays [102]. 2) Use a high-resolution TDC instead of a bang-band PD. If frac-N quantization noise spans many TDC thresholds, the effective phase detector gain will be constant, determined by the slope of a straight line fit through the staircase characteristic. Further, the phase detector's own quantization noise will be smaller and contribute less jitter at the loop output. 3) Use a linear charge pump-based phase detector, whose gain is always constant but which injects thermal noise into the loop instead of quantization noise—the two can be comparable, as we show in Section 3.4.

When designing for low output jitter a charge pump phase detector, in spite of its known imperfections [65, Ch. 9], should not be dismissed out-of-hand; of the three remedies listed, it might well involve the least design effort.

3.9 Design Guidelines

As argued in Sec.3.4, a bang-bang PLL performs equally well as a charge-pump PLL for integer-N frequency synthesizers. It may be preferred for its simpler circuits. The design of bang-bang PLL condenses to the choice of loop filter gain β because the reference frequency, divider ratio and phase noise of the chosen VCO are usually already fixed.

(3.25) reveals that the optimal β_{opt} which minimizes the output jitter is a noise-independent parameter. It remains a constant regardless of the change of reference phase noise. The VCO gain K_V is usually set by the C-V relationship of a MOS varactor. This is only determined by the process technology and has minimum dependence on temperature as shown by [47, Eq. (15)]. K_W does depend on absolute temperature, but a temperature variation of $[-20°\text{C}, +120°\text{C}]$ only changes β_{opt} by $\pm 10\%$ from its nominal at 27°C. Fig.3.16(b) shows a relatively large tolerance to β and 10% change does not raise the output jitter significantly. Thus, adaptation is not necessarily needed to optimize jitter.

Whereas minimizing output jitter in a bang-bang PLL requires β_{opt}, the out-of-band phase noise roll-off improves monotonically with smaller β (3.26). So for applications where both design requirements are to be met, one chooses the smallest β so long as the output jitter remains within specifications.

In fractional-N PLLs, although the divider phase noise lowers K_{BPD} and worsens the output jitter, β_{opt} that minimizes the output jitter will remain unchanged, (3.25). Following the same optimization as above, one can estimate if a bang-bang fractional-N PLL without divider noise cancellation is sufficient. If not, noise cancellations employing DTC/TDC may be needed. Or if feasible, a charge-pump may prove to be the quickest solution.

3.10 Conclusion

We have developed a simple and accurate design-oriented analysis for phase locked loops that use a bang-bang phase detector. An effective linearized gain of this phase detector is interpreted in signal space and a closed-form analytical expression is found. Since each

noise source's contribution at the loop's output can now be assessed through linear feedback analysis, it is possible to optimize loop design.

Compared to an analog linear PLL that would typically use a charge-pump, a bang-bang PLL may operate at higher frequencies, is easier to design and scales readily to advanced CMOS technologies. The analysis presented here helps to dispel two misconceptions. First, all else being equal a PLL with a bang-bang phase detector can be designed for the same output jitter as a loop with a charge pump. Second, the bang-bang PLL can be designed *a priori* for optimum jitter without the need for real-time adaptation during operation.

The frequency-domain analysis presented in this paper should enable better trade-offs and shorter design time compared to an excessive reliance on slow time-domain simulations.

3.11 Appendix

3.11.1 Fractional-N Divider Phase Noise

The divider phase error sampled by the phase detector at every reference cycle is the sum of the phase error at the end of previous cycle and the phase change through the current cycle caused by the fluctuating integer divider ratio around the average, N' (see Fig.15). Thus,

$$\begin{aligned}\phi_N[kT_{ref}] &= \phi_N[(k-1)T_{ref}] + \frac{\omega_{out}T_{ref}}{N_k} - \frac{\omega_{out}T_{ref}}{N'} \\ &= \phi_N[(k-1)T_{ref}] + \frac{2\pi}{N'} \cdot \Delta n_k \end{aligned} \tag{3.34}$$

where $\Delta n_k = N_k - N'$ represents the instantaneous departure from the desired non-integer divider ratio and from Fig.3.17, $\omega_{out}T_{ref} = 2\pi N'$. Then, following the analysis in [86],

$$\phi_N[kT_{ref}] = \frac{2\pi}{N'} \sum_{k=1}^{\infty} \Delta n_k \tag{3.35}$$

(3.35) represents integration in the discrete time domain and in the frequency domain is equivalent to

$$\Phi_N(z) = \frac{2\pi}{N'} \cdot \frac{S_{\Delta n}}{1 - z^{-1}} \tag{3.36}$$

Δn is the result of a white quantization noise after being shaped by the Σ-Δ modulator. Since the divider ratio is updated every reference cycle, $z = e^{j2\pi f T_{ref}}$. Depending on the order m of the modulator, its PSD is characterized as

$$S_{\Delta n} = \frac{1}{12} \cdot \frac{2}{f_{ref}} \left|(1 - z^{-1})^m\right|^2 \tag{3.37}$$

Noting that $\langle \phi_N^2 \rangle$ finally sets the detector gain K_{BPD}, we can obtain the following closed-form expressions using (3.36) and (3.37):

$$\begin{aligned}
\langle \phi_N^2 \rangle &= \left(\frac{2\pi}{N'}\right)^2 \frac{1}{12} \cdot \frac{2}{f_{ref}} \int_0^{\frac{f_{ref}}{2}} \left|\left(1 - e^{-j2\pi f T}\right)^{m-1}\right|^2 \mathrm{d}f \\
&= \left(\frac{2\pi}{N'}\right)^2 \frac{1}{12} \cdot \frac{2}{f_{ref}} \int_0^{\frac{f_{ref}}{2}} \left(2\sin(\pi f T_{ref})\right)^{2(m-1)} \mathrm{d}f \\
&= \begin{cases}
\left(\frac{2\pi}{N'}\right)^2 \times \frac{1}{12} & \text{for} \quad m = 1 \\
\left(\frac{2\pi}{N'}\right)^2 \times \frac{1}{12} \times 2 & \text{for} \quad m = 2 \\
\left(\frac{2\pi}{N'}\right)^2 \times \frac{1}{12} \times 6 & \text{for} \quad m = 3
\end{cases}
\end{aligned} \tag{3.38}$$

The order of quantization noise shaping is reduced by one due to the integration by the divider, as seen in (3.35).

CHAPTER 4

A 2.5GS/s 10bit 65mW 8-Channel Interleaving ADC in 28nm CMOS FD-SOI

4.1 Introduction

Wideband direct sampling receiver for communication systems has been of interest for many years. Unlike conventional narrowband receiver, RF signal is directly sampled and digitized by a high-resolution high-speed analog-to-digital converter(ADC). In this architecture, ADC captures the entire spectrum of interest, whereas all down-conversions and demodulations are done in digital domain. Compared to conventional narrow-band receivers, the full-band capture offers many advantages, such as wide bandwidth, high data rate, high agility, simultaneous reception of many channels, fast channel switching, low system complexity, and low cost. In recent years, the emergence of multi-GS/s high-resolution ADCs has enabled GHz-wide full-band capture in cable, satellite and networking applications. Further reduction of ADC power dissipation will allow it to expand into new application areas. We propose to build a full band capture ADC that is amplifier-less and purely comparator based to achieve better power-speed tradeoff.

4.2 Current State of Art

Most high-speed high-resolution ADCs can be divided into three categories: pipeline ADCs, interleaved successive approximation register(SAR) ADCs and flash ADCs. Sigma-delta ADCs are not suitable for high speed application because of limitation in oversampling ratio.

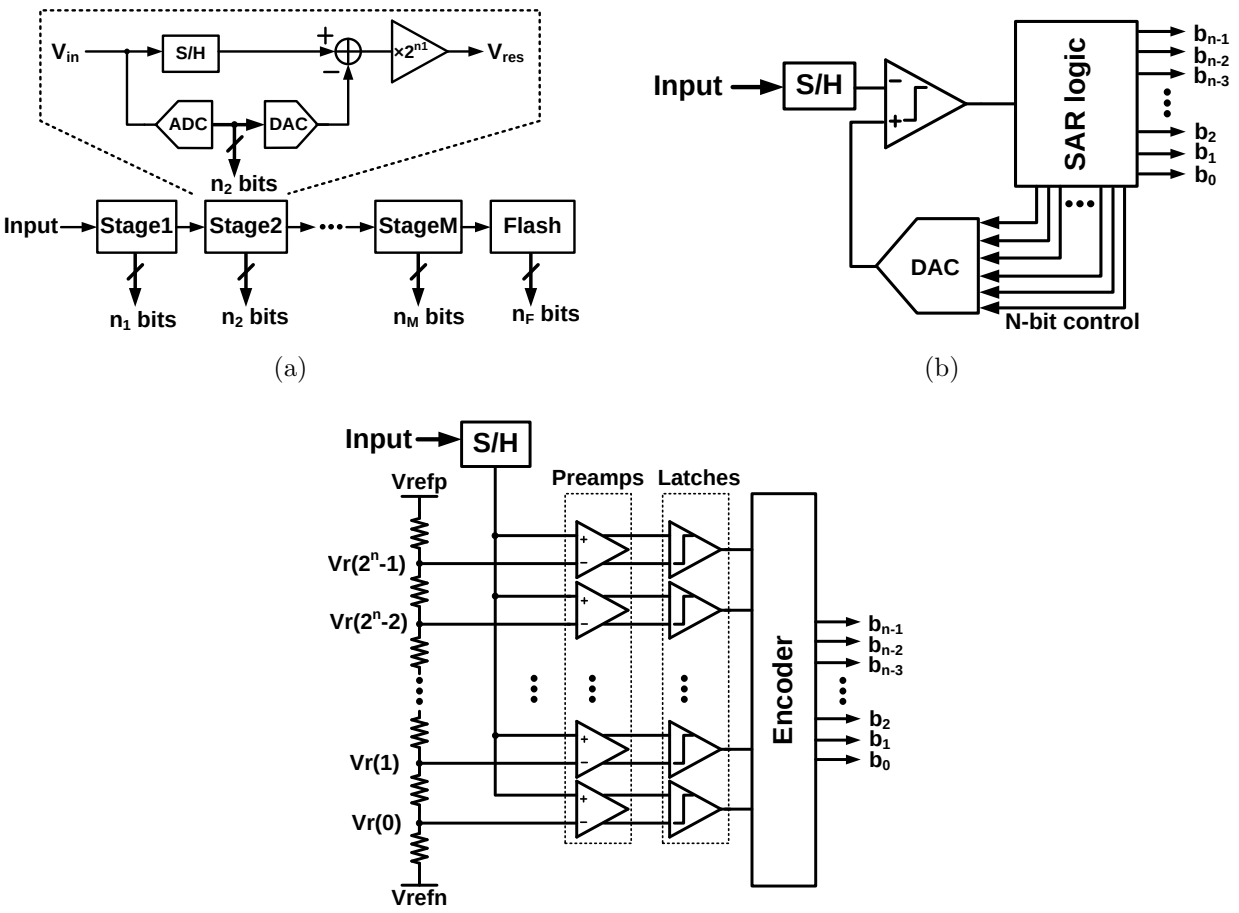


Figure 4.1: (a)Typical pipeline ADC architecture;(b)SAR ADC architecture; (c)one stage full-flash ADC architecture.

Pipeline ADCs have been one of the most popular choices for high-speed and high-resolution due to simple configuration and relatively high throughput. A conventional pipeline architecture is shown in Fig.4.1(a). The most significant power consumption comes from the linear amplifier that provides precise gain to amplify the residue voltage. Precise gain is usually implemented with a high-gain high-bandwidth opamp with feedback network. Various digital calibrations have been applied to compensate analog error introduced by amplifiers in pipeline ADCs, which has enabled faster pipeline ADCs with less power consumed by amplifiers. [103] finds nonlinear error of the opamp and applies its inverse function to correct this error. [104] models incomplete settling between stages as inter-stage-

interference(ISI) and uses FIR filters to compensate ISI. A single channel pipeline ADC has been able to achieve multiple GS/s 10bit resolution with power consumption around 240mW [105].

The successive approximation register(SAR) ADC implements binary search algorithm with one comparator, one DAC of the same resolution as ADC, a sample-and-hold(S/H) circuit and digital control logic as shown in Fig.1(b). SAR ADC usually uses only one comparator, but requires several iterations to finish conversion process. Thus, it is most efficient operating at MS/s and its application in multiple GS/s is limited, especially with total capacitance of DAC exponentially growing with resolution. To increase speed, time-interleaving is required. However, time-interleaving has issues itself due to mismatch among sub-channels. Calibrations correcting interleaving errors in time-interleaving ADCs has been extensively studied over the past ten years, but still, calibrations demands significant hardware and power overhead, especially for those operating above multiple GS/s range. Time-interleaving SAR ADCs with 5-7bit resolution have been widely explored [106, 107, 108, 109, 110, 111, 112, 113] because both time-interleaving and SAR ADCs are efficient for low resolution applications. The calibration becomes more sophisticated for interleaving ADCs to achieve 10 bit resolutions [114, 115, 116].

A full flash ADC directly compares input signal with transition points between adjacent quantization intervals as shown in Fig.1(c). It requires 2^N comparators for N-bit resolution and each comparator needs to have the same resolution as the overall ADC. As small size latches have large input referred offsets, preamplifiers are usually required to suppress offsets [117, 118], which become the dominant source of power consumption. The difficulty of implementing low-power high-resolution comparators has made flash ADCs unfavorable for high speed applications since they require large numbers of comparators. It is usually used as the first stage sub-ADC that samples and quantizes the input at full speed followed by more sub-interleaving ADCs [115, 116].

We propose to build a 2.5GS/s 10bit 8-channel interleaving SAR ADC with under 100mW power consumption. The aim of the work is to demonstrate a simple solution without intensive digital calibrations for multi-GS/s ADCs with moderate resolutions.

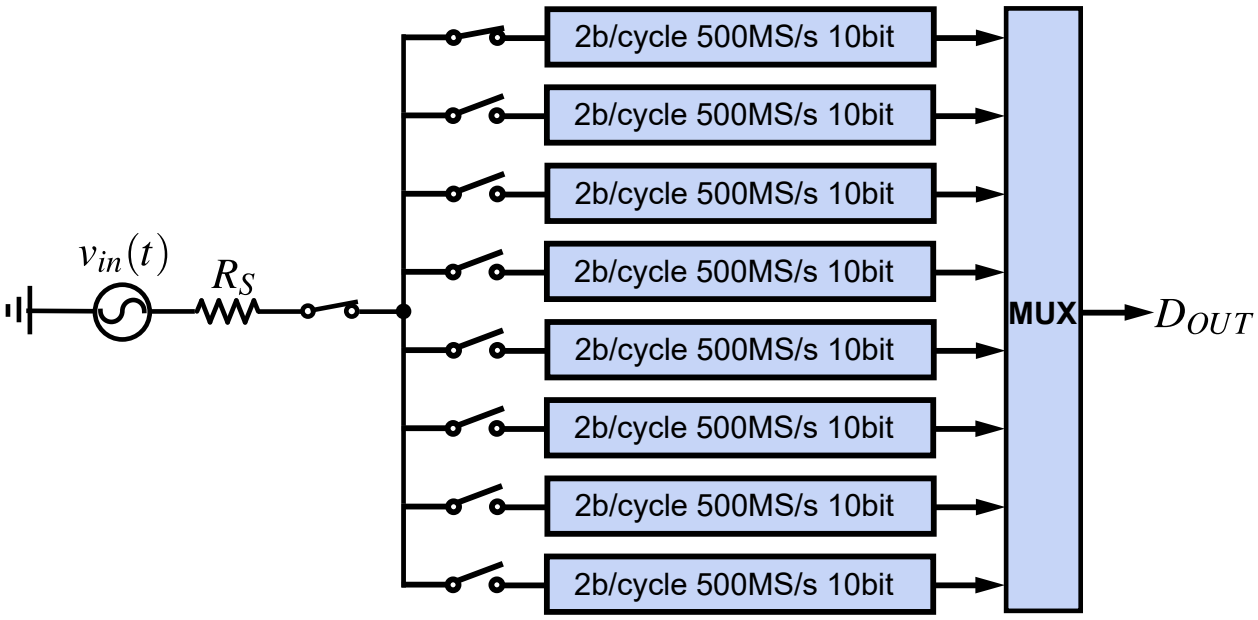


Figure 4.2: System architecture.

The key techniques to optimize the design trade-offs include:

- Fully passive master/slave bottom plate sampling scheme to suppress timing skew. This involves global optimization of single channel distortion, residual timing skew and bandwidth mismatch among channels.
- 2b/cycle sub-channel SAR ADC with calibrated comparators for faster conversion and smaller total capacitance in capacitor array.
- Use calibrated comparators to eliminate the need for preamplifiers.

Rather than intensive digital calibrations, we explore fundamental circuit techniques to address degradations arising from mismatches among interleaving channels, namely gain mismatch, offset mismatch, timing mismatch and bandwidth mismatch.

4.3 Challenges in Interleaving ADCs

Conceptually, by interleaving N ADCs each operating at a sampling rate of f_s/N with M bit resolution, the overall interleaved ADC can potentially achieve a sampling rate of f_s with

M bit resolution. Principles of interleaving is best understood in frequency domain. Since each sub-channel ADC sub-samples the input signal that creates multiple images below the Nyquist frequency, successfully reconstructing output signal demands correct phase and gain alignment among sub-channels to cancel unwanted images.

$$\begin{aligned} Y_{CH_0}(j\omega) &= X_{in}(j\omega) * \sum_{n=-\infty}^{+\infty} \delta(\omega - n \times \frac{\omega_s}{N}) \\ Y_{CH_1}(j\omega) &= X_{in}(j\omega) * \sum_{n=-\infty}^{+\infty} \delta(\omega - n \times \frac{\omega_s}{N}) e^{-j\omega T_s} \\ &\vdots \\ Y_{CH_{N-1}}(j\omega) &= X_{in}(j\omega) * \sum_{n=-\infty}^{+\infty} \delta(\omega - n \times \frac{\omega_s}{N}) e^{-j\omega[(N-1)T_s]} \end{aligned} \tag{4.1}$$

The overall output is the summation of all sub-channels

$$\begin{aligned} Y_{out}(j\omega) &= \sum_{i=0}^{N-1} Y_{CH_i}(j\omega) \\ &= X_{in}(j\omega) * \sum_{n=-\infty}^{+\infty} \delta(\omega - n \times \omega_s) \end{aligned} \tag{4.2}$$

The interleaved ADC operates as if it was a single channel ADC operating at full sampling rate of f_s. Yet, any interleaved ADC suffers from non-ideality factors that limit its overall resolution below than a single sub-channel ADC. These sources of degradation include gain mismatches, offset mismatches, timing skews and tracking bandwidth mismatches. Including these non-ideality factors, output of each sub-channel is modified to

$$\begin{aligned} Y_{CH_0}(j\omega) &= [X_{in}(j\omega) + v_{OS_1}] \cdot \frac{1}{1 + j\frac{\omega}{\omega_{TH_1}}} * G_1 \sum_{n=-\infty}^{+\infty} \delta(\omega - n \times \frac{\omega_s}{N}) e^{-j\omega\Delta t_1} \\ Y_{CH_1}(j\omega) &= [X_{in}(j\omega) + v_{OS_2}] \cdot \frac{1}{1 + j\frac{\omega}{\omega_{TH_2}}} * G_2 \sum_{n=-\infty}^{+\infty} \delta(\omega - n \times \frac{\omega_s}{N}) e^{-j\omega(T_s + \Delta t_2)} \\ &\vdots \\ Y_{CH_{N-1}}(j\omega) &= [X_{in}(j\omega) + v_{OS_{N-1}}] \cdot \frac{1}{1 + j\frac{\omega}{\omega_{TH_{N-1}}}} \\ &\quad * G_{N-1} \sum_{n=-\infty}^{+\infty} \delta(\omega - n \times \frac{\omega_s}{N}) e^{-j\omega[(N-1)T_s + \Delta t_{N-1}]} \end{aligned} \tag{4.3}$$

where v_{OS_i} represents offset introduced in each channel, ω_{TH_i} represents the tracking bandwidth of each channel, G_i represents the gain of each channel and Δt_i represents the sampling instant shifted by timing skews. The overall output then contains unwanted interleaving spurs, which do not have closed form expressions, on top of wanted signal spectrum that needs to be reconstructed. Since these spurs fall below Nyquist frequency, they must be suppressed or corrected.

$$\begin{aligned} Y_{out}(j\omega) &= \sum_{i=0}^{N-1} Y_{CH_i}(j\omega) \\ &= X_{in}(j\omega) * \sum_{n=-\infty}^{+\infty} \delta(\omega - n \times \omega_s) + \text{Interleaving spurs\&images} \end{aligned} \tag{4.4}$$

In the following subsections, we examine impact of each non-ideality factor.

4.3.1 Offset mismatch

Offset in a single-channel ADC introduces only a DC component on the output spectrum. This does not interfere with signal reconstruction and can be ignored. However, in interleaved ADCs, because of sampling action in each sub-channel, DC offset in sub-channel ADCs is converted to spurs located at $\frac{f_s}{N}$, $\frac{2f_s}{N}$, $\frac{3f_s}{N}$ ... ...

These interleaving spurs are *independent* of input signal amplitude and frequency. Their amplitudes are determined by amount of offset mismatch and frequencies are fixed at sub-sampling frequencies. As they are largely static tones on spectrum, back-end digital signal processing can relatively easily remove these spurs. On the other hand, designing each sub-channel with minimum offset is another viable solution as offset is already well understood [22].

4.3.2 Gain mismatch

When two ADCs have two different full scale voltages, they will display different gains from the input to output, albeit they are perfectly linear by each own. This mismatch in gain leads to incomplete cancellation of images, creating images at $\frac{f_s}{N} \pm f_{in}$, $\frac{2f_s}{N} \pm f_{in}$, $\frac{3f_s}{N} \pm f_{in}$ ...

... These images are related to input amplitude and frequency.

4.3.3 Timing skew

Timing skew occurs when sub-channel ADCs are not sampling exactly at time interval of $T_s = \frac{1}{f_s}$. Timing skews leads to incomplete cancellation of images and spurs, resulting in both images at $\frac{f_s}{N} \pm f_{in}$, $\frac{2f_s}{N} \pm f_{in}$, $\frac{3f_s}{N} \pm f_{in}$ ... ... and fixed frequency spurs at $\frac{f_s}{N}$, $\frac{2f_s}{N}$, $\frac{3f_s}{N}$ ... ...

4.3.4 Bandwidth mismatch

Practically, a T/H circuit can be approximated as a single pole transfer function with unity DC gain.

$$\begin{aligned} H(j\omega_{in}) &= \frac{1}{1+\frac{j\omega_{in}}{\omega_0+\Delta\omega_0}} \\ &\approx \frac{1}{\exp\left(\frac{\omega_{in}}{\omega_0+\Delta\omega_0}\right)} \\ &= \exp\Big(-j\frac{\omega_{in}}{\omega_0+\Delta\omega_0}\Big) \\ &= \exp\Big[-j\frac{\omega_{in}}{\omega_0}(1-\frac{\Delta\omega_0}{\omega_0})\Big] \\ &= \underbrace{\exp\Big(-j\frac{\omega_{in}}{\omega_0}\Big)}_{\text{nominal}} \cdot \underbrace{\exp\Big(-j\omega_{in}\cdot\frac{\Delta\omega_0}{\omega_0^2}\Big)}_{\text{equivalent skew}} \end{aligned} \tag{4.5}$$

The additional equivalent skew is

$$\Delta t_{BW} = \frac{\Delta\omega_0}{\omega_0^2} = \underbrace{\frac{\Delta\omega_0}{\omega_0}}_{\text{percentage}} \times \underbrace{\frac{1}{\omega_0}}_{\text{RC constant}} \tag{4.6}$$

Small bandwidth mismatch is equivalent to timing skew. Its degradation on ADC resolution reaches maximum at Nyquist input frequency with no impact on performance at DC input.

Since bandwidth mismatch is usually caused by mismatches in resistors implemented as transistors, it is beneficial to further simplify (4.6) (capacitor matching is usually not a

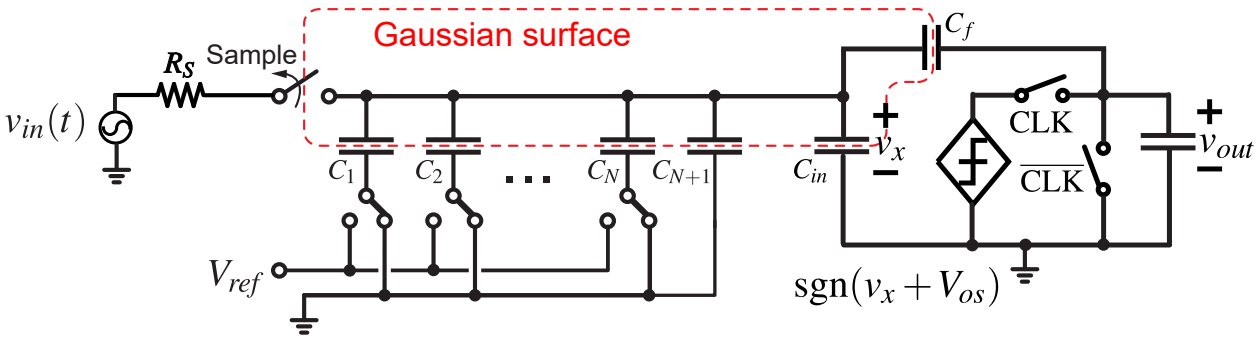


Figure 4.3: SAR ADC with top-plate sampling scheme.

limitation as it relies on geometries).

$$\begin{aligned} \Delta t_{BW} &= \frac{\Delta\omega_0}{\omega_0} \times \frac{1}{\omega_0} \\ &= \frac{\Delta R}{R} \times RC \end{aligned} \tag{4.7}$$

(4.6) offers practical guidance to suppress the impact of bandwidth mismatch. Besides minimizing the mismatch in percentage, it is also essential to maximize the tracking bandwidth. Higher bandwidths reduce the equivalent skew with the same bandwidth mismatch percentage.

4.4 Principles of Single Channel SAR ADC

4.4.1 Top-Plate vs. Bottom-Plate Sampling

Before delving into detailed construction of the ADC, we first investigate the first principle of SAR ADCs. Shown in Fig.4.3 is a singled-ended SAR ADC utilizing top-plate sampling scheme. C_1-C_{N+1} are capacitors in the capacitive DAC(CDAC), C_{in} represents the input capacitor of the comparator and C_f represents the feedback capacitor of the comparator. C_1-C_{N+1} are linear capacitors, C_{in} is voltage dependent as it is dominated by the MOS cap of a transistor, C_f comes from transistor parasitics and creates a feedback path that results in 'kick-back'. The controlled source $\text{sgn}(v_x + V_{OS})$ represents the decision function of the comparator, along with the loading capacitor which is reset by CLK after each decision.

During sampling phase, the input is connected to the top plates of the capacitors with

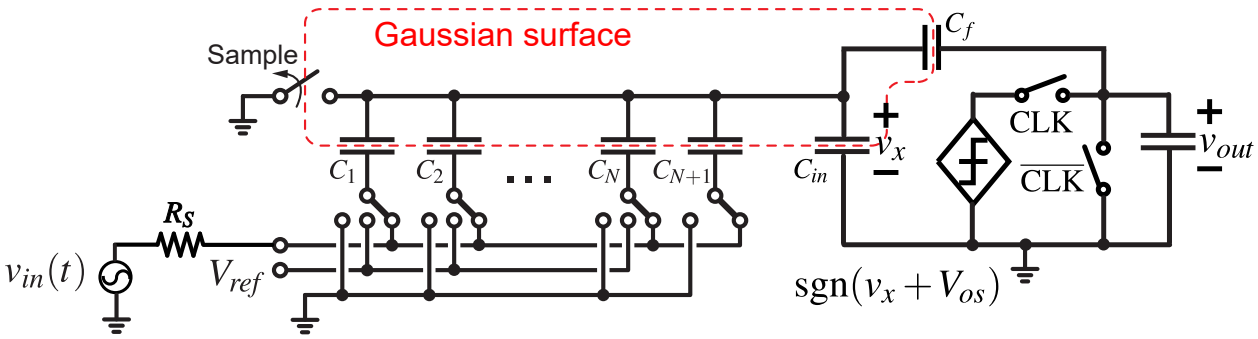


Figure 4.4: SAR ADC with bottom-plate sampling scheme.

the bottom plates grounded, creating a voltage drop of $v_{in}(0)$ at the end of sampling phase. Assume the switches are ideal, $v_{in}(0)$ does not contain any distortion component during the sampling phase.

Through the conversion phase, once the sampling switch is turned off, the total charge contained in DAC capacitors C_1-C_N, C_{in} and C_f, which forms a Gaussian surface, is preserved and solely determined by the sampling phase. Bottom plates are connected to either V_{ref} or ground depending on the decision feedback from the comparator, driving the residual voltage v_x at the comparator input to zero (within 1LSB) at the end of conversion. Using charge conservation within the Gaussian surface, we have the relationship below

$$\begin{aligned} v_{in}(0)(C + C_{in} + C_f) &= \sum_{1}^{N} C_k \left(v_x - b_k V_{ref}\right) + v_x(C_{in} + C_f) \\ \Rightarrow D_{out} &= \sum_{1}^{N} \frac{b_k}{2^{N-1}} \\ &= \underbrace{\frac{C + C_{in} + C_f}{C} \frac{v_{in}(0)}{V_{ref}}}_{\text{Gain}} - \underbrace{\frac{C + C_{in} + C_f}{C} \frac{v_x}{V_{ref}}}_{\text{Quantization}} \end{aligned} \tag{4.8}$$

At the end of conversion phase, v_x is forced to be less than 1LSB by SAR algorithm. However, because C_{in} is voltage dependent and v_x has different values during sampling phase and conversion phase, the gain term $(C+C_{in}+C_f)/C$ is input dependent. The feedback coupling capacitor C_f is also voltage dependent, but its impact is usually less than C_{in}. This translates to distortions after reconstructing the digitized signal. Thus, a SAR ADC using top-plate sampling scheme is non-linear by its nature.

On the other hand, in a SAR ADC using bottom-plate sampling scheme as shown in Fig.4.4, the impact of C_{in} is removed. During sampling phase, bottom plates of capacitor in the CDAC are connected to the input while the top plates are all connected to ground. Again, the total charge in the Gaussian surface is defined by the sampling phase. During conversion phase, SAR algorithm switches the bottom plates of each capacitor in CDAC to either V_{ref} or ground, forcing the residual voltage v_x to approximately zero. With charge conservation, we have

$$\begin{aligned} v_{in}(0)C &= \sum_{1}^{N} C_k \left(v_x - b_k V_{ref}\right) + v_x(C_{in} + C_f) \\ \Rightarrow D_{out} &= \sum_{1}^{N} \frac{b_k}{2^{N-1}} \\ &= \underbrace{-\frac{v_{in}(0)}{V_{ref}}}_{\text{Gain}} + \underbrace{\frac{C + C_{in} + C_f}{C}\frac{v_x}{V_{ref}}}_{\text{Quantization}} \end{aligned} \tag{4.9}$$

Since C_{in},C_f do not carry charge before/after conversion, they will not affect the overall linearity of the ADC.

Comparing top-plate and bottom-plate sampling schemes in a SAR ADC, besides the benefit of lower distortion from track/hold, the more important benefit of bottom-plate sampling is that it removes the impact of nonlinear C_{in} from the conversion algorithm. For 10-12 bit applications, C_{in} significantly degrades the overall linearity and thus, bottom-plate sampling scheme should always be preferred in moderate to high resolution ADCs.

4.5 Fully Passive Track/Hold Circuits

4.5.1 Distortion in a simple passive R-C track/hold circuit

Besides conversion, the overall linearity of an ADC is as well limited by the track/hold circuit. The sampled voltage $v_{in}(0)$ should be linear to the analog input with distortion below design spec.

In a simple passive R-C track/hold circuit, there are three distortion sources degrading

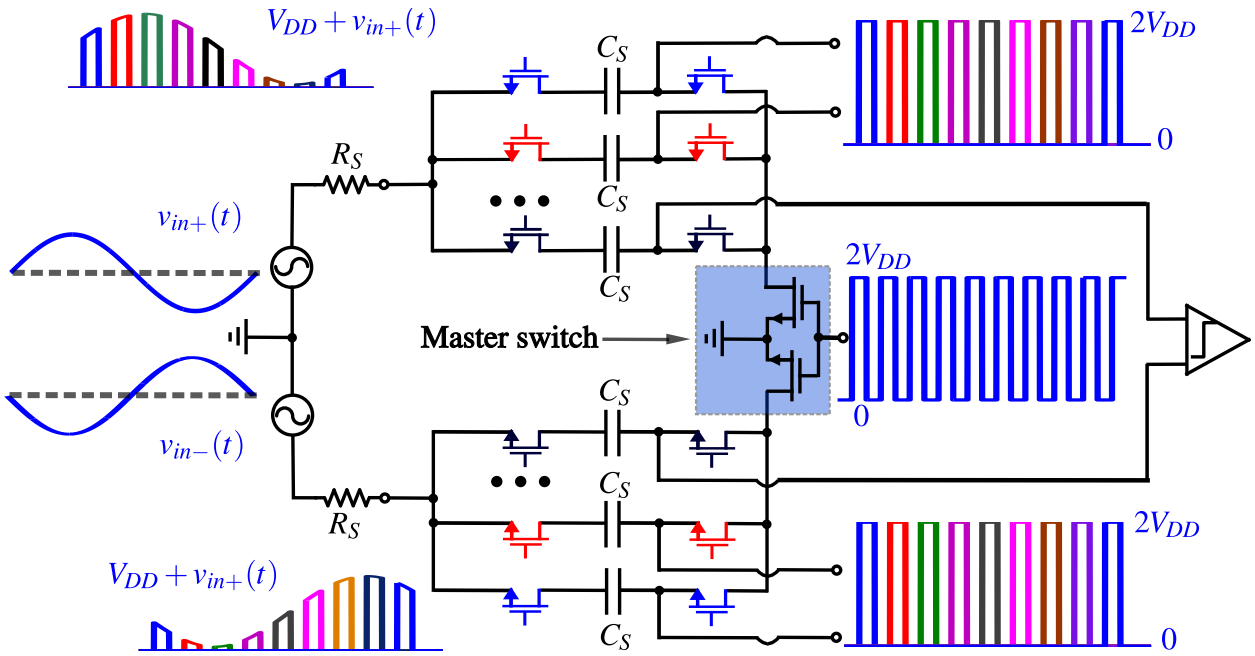


Figure 4.5: Passive master-slave sampling scheme in interleaving SAR ADCs.

linearity, the nonlinear resistance modulation from R_{ON}, signal dependent charge-injection and turn-off timing modulation due to finite slope of gate voltage [40].

- R_{ON} distortion increases with input frequency and is inversely proportional to the RC bandwidth.

$$HD3_{R_{on}} \approx \frac{n^2 C_2 \omega_0}{4\beta V_{ov}^3} A_0^2 = \frac{1}{4}\left(\frac{nA_0}{V_{ov}}\right)^2 \underbrace{\frac{C_2}{\beta V_{ov}}}_{\sim BW^{-1}} \omega_0 \tag{4.10}$$

- Charge injection distortion has a more complicated expression, but generally proportional to the transistor width and related to clock transition time τ.

$$HD3_{ci} = \frac{n^3}{48}\frac{A_0^2}{V_{DD}}\frac{C_G}{C_2}\frac{\beta}{C_1 \parallel C_2}\tau\left(\frac{V_{ov}}{V_{DD}}\frac{\beta V_{ov}}{C_1 \parallel C_2}\tau - 1\right)\exp\left(-\frac{1}{2}\frac{V_{ov}}{V_{DD}}\frac{\beta V_{ov}}{C_1 \parallel C_2}\tau\right) \tag{4.11}$$

- Turn-off timing depends only on the transition time τ

$$HD3_{tim} \approx \frac{n^2}{8}\left(\frac{A_0}{V_{DD}}\right)^2 (\omega_0 \tau)^2 \tag{4.12}$$

A_0 represents the input amplitude, n is the slope factor in a certain process, usually $1.2 \sim 1.4$, ω_0 is the input frequency, C_2 is the load capacitance and C_G is the total gate capacitance of the transistor.

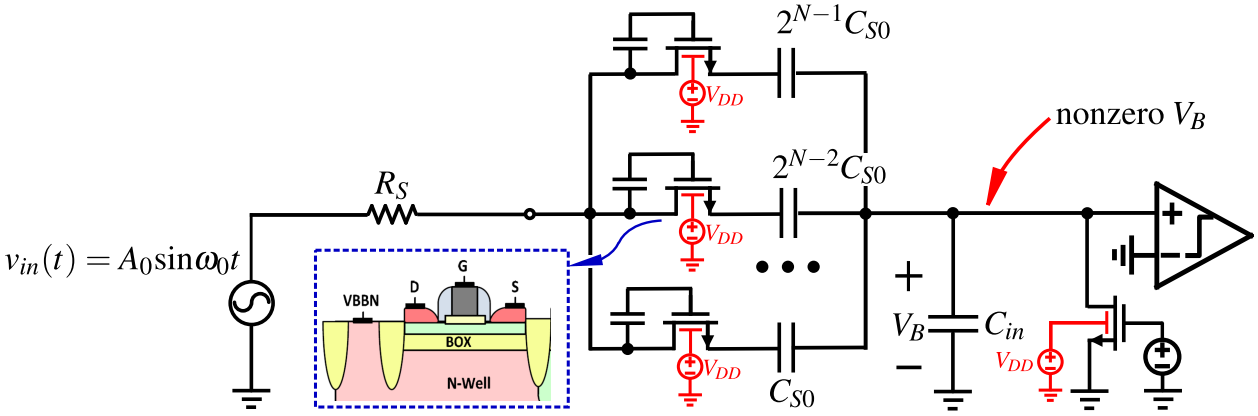


Figure 4.6: SAR ADC with bottom-plate sampling scheme.

For all practical purposes, turn-off timing is negligible in modern CMOS processes. Thus the optimization of a track/hold condenses to choosing the correct transistor width to balance the contribution from R_{on} and charge-injection modulations.

4.5.2 Distortion in bottom-plate track/hold in a SAR ADC

For a SAR ADC with passive bottom-plate track/hold scheme, the detailed circuits in a single-ended version during sampling phase is shown in Fig.4.6. The 50Ω input source is connected to top plates of CDAC capacitors while the bottom plates are connected to a DC voltage. The top plate switches are bootstrapped and the gate voltage during sampling phase is $v_{in}(t)+V_{DD}$. This suppresses R_{on} distortion with a constant V_{GS}. The bottom plate switches are bootstrapped with gate voltage of $2V_{DD}$ during sampling phase to maximize the tracking bandwidth. The bottom-plate switch is turned off ahead of top-plate switches to suppress charge-injection modulation from the top-plate switches.

Beyond the techniques above, distortions still limit the linearity of the track/hold circuits [40].

- Residual nonlinear R_{ON} because body still modulates R_{ON}. In FD-SOI process, this is especially ominous because body has a strong ability controlling threshold voltage of a transistor.

$$\mathrm{HD3}_{R_{ON1}} = \frac{1}{4}\left[\frac{(n-1)A_0}{V_{OV1}}\right]^2 \frac{C_S}{\beta V_{OV1}} \cdot \omega_0 \tag{4.13}$$

- Since V_B is a finite voltage and proportional to the input, both R_{ON} modulation and charge-injection modulation from bottom-plate switch contribute to distortions.
  - R_{ON} modulation depends on both the signal swing across the bottom switch ΔA_0 and its size.
$$\mathrm{HD3}_{R_{ON2}} = \frac{1}{4}\left(\frac{n\Delta A_0}{V_{OV2}}\right)^2 \frac{C_S}{\beta V_{OV2}} \cdot \omega_0$$
  - Similarly, charge-injection modulation depends on both signal swing and transistor size.

These observations lead to the optimization guidelines for the bottom-plate track/hold circuit. For the top-plate switch, its major contribution to distortion to the sampled output is R_{ON} distortion. Increasing its size reduces this distortion, but its size cannot be arbitrarily big as extra power needs to be consumed in the bootstrapping circuit to drive its gate. Its junction capacitors, which is nonlinear by itself, also limits its size. Another limitation on the size of the top-plate switch is that the input feedthrough interrupts conversion when the switch is turned off (will be discussed in subsequent section). For the bottom-plate switch, R_{ON} distortion monotonically decreases with larger transistor size, which simultaneously leads to larger RC bandwidth and smaller signal swing across bottom switch. Charge injection also decreases with larger transistor size as the benefit of reduced signal swing dominates over impact of larger transistor size. Another limiting factor comes from the nonlinear junction capacitor from a MOSFET, which will degrade overall linearity significantly if it is sized too large.

4.5.3 Body biasing of the top-plate switch

As discussed previously, boot-strapping top-plate switch does not fully remove R_{ON} modulation due to body effect. Since body is also used as 'back gate' with strong ability of controlling threshold voltage, body effect in FD-SOI is even stronger than standard CMOS, resulting in a slope factor of $n \sim 1.5$ [119]. This limitation can be further eliminated if body of the transistor is connected to the source as shown in Fig.4.7. However, this introduces another issue, the input feedthrough.

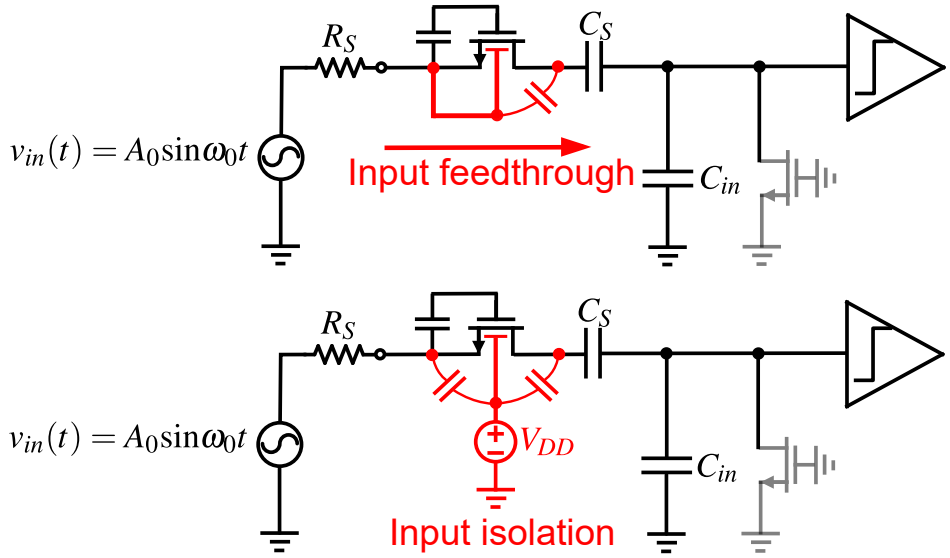


Figure 4.7: SAR ADC with bottom-plate sampling scheme.

Tying the body and sources creates a capacitive path for the input signal to couple to comparator input even when the gate voltage is zero. This can distort the held voltage, lead the comparator to make wrong decisions and interrupt the conversion operation. With Nyquist input of 2GHz and $C_S = 300fF$, $W/L = 30\mu$m/28nm, input feedthrough can create a 5mV ripple at comparator input, almost as 3.5 times as 1LSB.

Tying body to a fixed DC voltage provides better isolation between the input source and the comparator. In this design, body is connected to V_{DD} to minimize the on-resistance R_{ON} and thus the sampling bandwidth. With the same input frequency and transistor size, tying body to V_{DD} reduces the bottom-plate fluctuation to 0.25mV, around 20 times improvement compared to the case when body is connected to the input signal.

4.5.4 Subchannel track/hold in master/slave sampling scheme

With a complete understanding of passive track/hold circuit, we investigate the design optimization of the bottom-plate track/hold in a single subchannel SAR ADC Fig.4.9, which will be used in interleaving channels with master/slave sampling scheme. SW_M is the master switch driven by the full rate clock and SW_S is the slave switch driven by the divided clock. Master switch is first turned off, then slave switch and finally top-plate switch. In

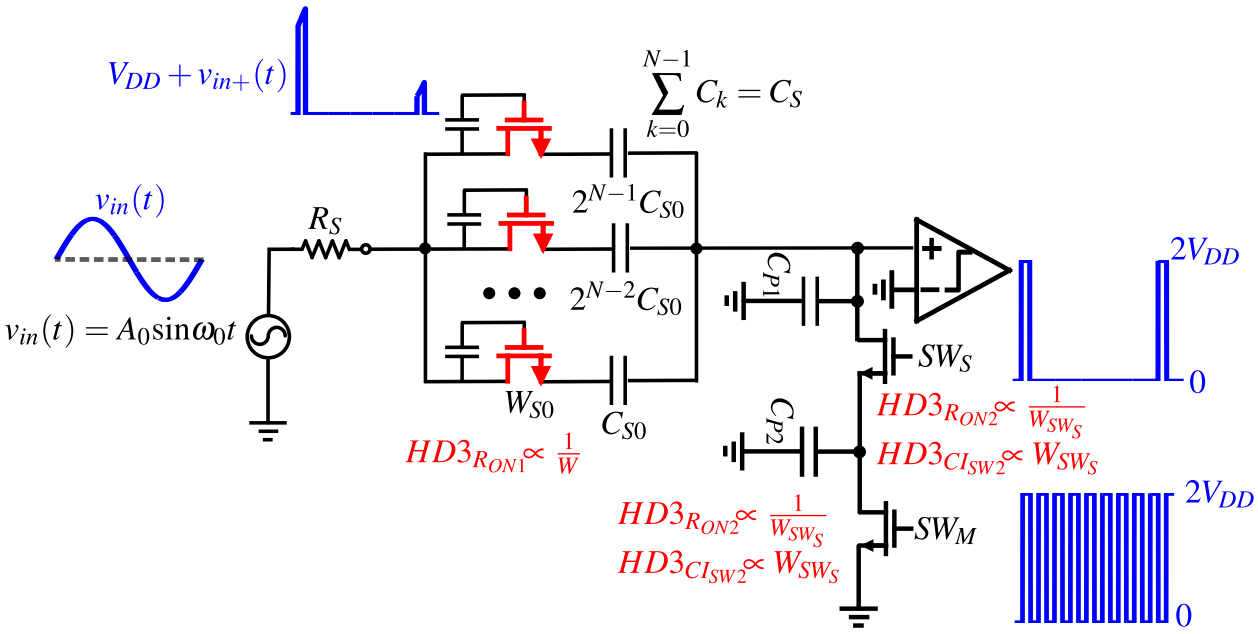


Figure 4.8: SAR ADC with bottom-plate sampling scheme.

this arrangement, the sampling instant will be solely determined by the master switch alone. The top-plate switches are driven by driven by boot-strapped voltage from the input. The bottom switches are both driven by voltages boot-strapped to $2V_{DD}$.

With three transistor in series connection, the residual signal swing across SW_M and SW_S becomes larger than the case previously discussed, resulting in worse distortions. This is the trade-off between suppressing interleaving spurs and achieving high linearity of a single channel. The distortion sources include

- R_{ON} distortion from top-plate switch $\propto W$;
- R_{ON} distortion from slave switch $\propto \frac{1}{W_S}$;
- Charge injection from slave switch $\propto W_S$;
- R_{ON} distortion from master switch $\propto \frac{1}{W_M}$;
- Charge injection from master switch $\propto W_M$.

Except higher distortions, another limit of track/hold with master/slave switch comes

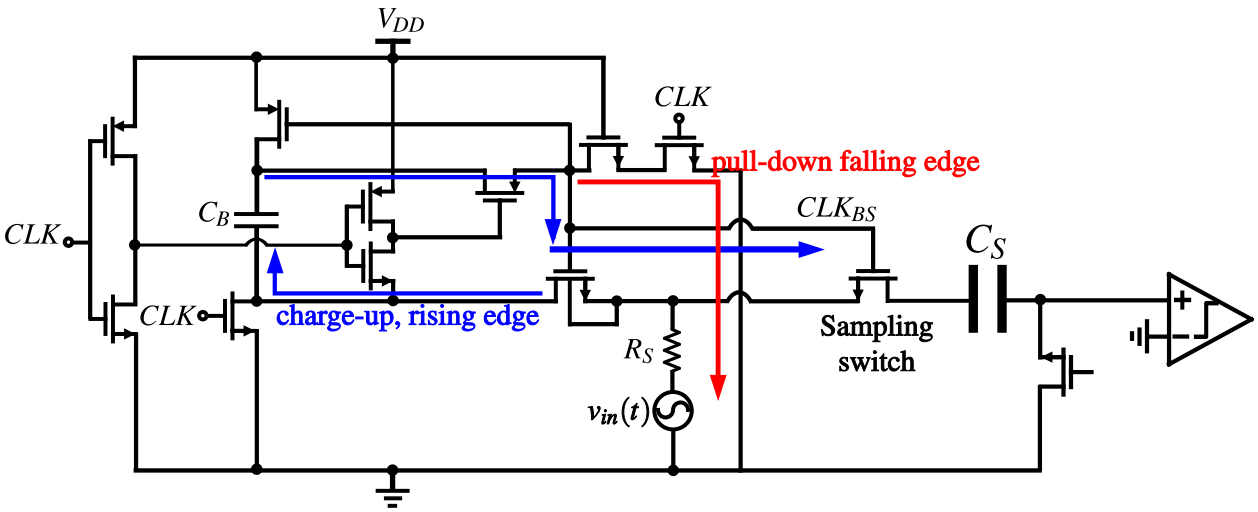


Figure 4.9: Boot-strapping circuitry.

from the total resistance of these series switches.

$$\text{BW} = \frac{1}{2\pi RC} = \frac{1}{2\pi \times 100\Omega \times 300\text{fF}} = 5.5\text{GHz} \tag{4.14}$$

This has limited the sampling bandwidth to 5.5GHz. High sampling bandwidth is preferred for complete settling that the sampled voltage on the capacitor becomes equal to the input voltage. Reduced sampling bandwidth does not degrade linearity, but poses difficulty in matching different subchannels. As we will show later, bandwidth mismatch is directly related to sampling bandwidth of each subchannel and high sampling bandwidth suppresses bandwidth mismatch.

Over 70dB SFDR is achieved through Nyquist input frequency in schematic level simulations. After extraction, SFDR degrades by ∼10dB at Nyquist frequency mainly because of resistive interconnect, resulting in larger signal swings across bottom plate switches, which further creates linearity degradations.

4.6 2bit/cycle SAR ADC

1bit/cycle SAR ADC utilizes one DAC and one comparator to make one decision bit per cycle. 2bit/cycle SAR ADC uses a flash ADC to accelerate the conversion [120, 121]. With 3 DACs generating 3 three threshold voltages compared with the sampled input voltage by

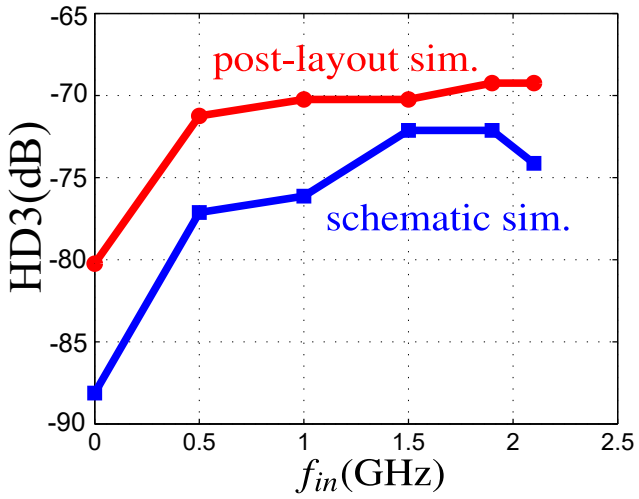


Figure 4.10: Schematic and layout extracted SFDR simulation of track/hold circuit with master/slave switch. Input amplitude $A_0 = 0.8$V, top-plate switch size $W/L = 30\mu$m/28nm, slave switch $W_S = 8\mu$m/28nm, master switch $W_M = 18\mu$m/28nm, total sampling capacitor $C_S = 300$fF.

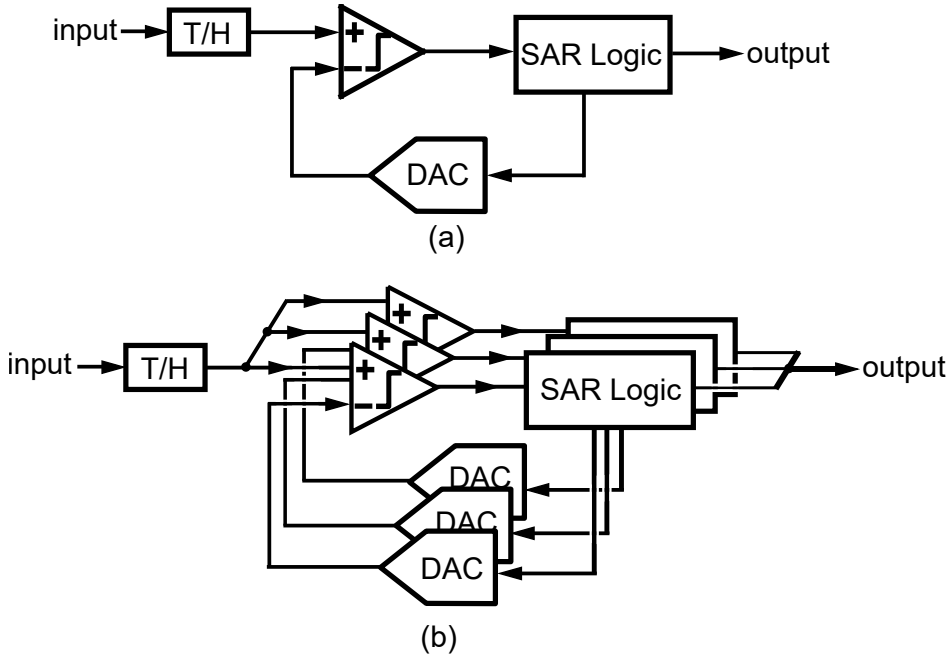


Figure 4.11: (a) 1bit/cycle SAR ADC; (b) 2bit/cycle SAR ADC.

3 comparators, the ADC is able to detect 2 bits every cycle. The 2 times conversion speed improvement comes at the cost of 3 times the hardware cost as shown in Fig.4.11.

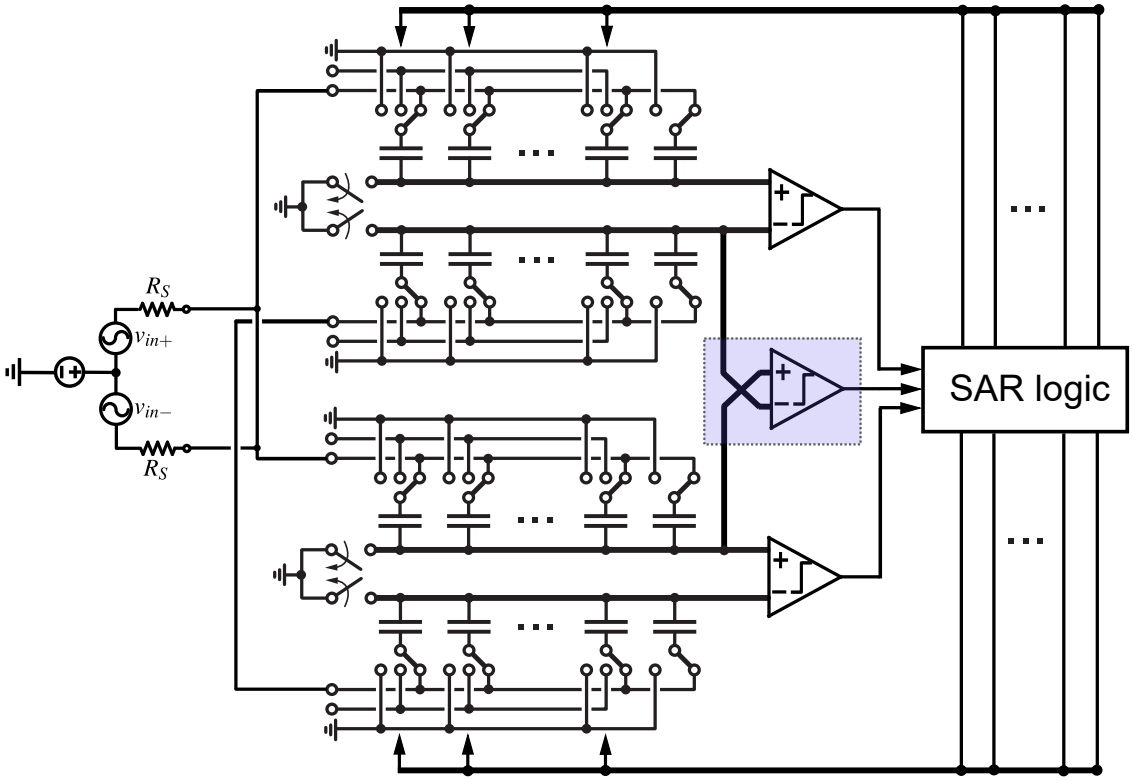


Figure 4.12: Differential 2bit/cycle SAR ADC using interpolation to remove one capacitive DAC.

Now we further investigate the design of a charge-redistribution 2bit/cycle SAR ADC and demonstrate that the speed improvement can be achieved with less hardware overhead. In a fully differential configuration with capacitive DACs, one array of capacitors can be removed with interpolation, through which one differential voltage is generated from the other two capacitive DACs as shown in Fig.4.12. Removal of one capacitive DAC also reduces the digital logic gates that would be needed to drive the DAC otherwise. Instead of comparing the input signal with one threshold voltage per comparison cycle in a 1bit/cycle SAR ADC, now the input voltage is compared with three threshold voltages per comparison cycle, enabling the 2bit/cycle SAR ADC to resolve 2bits per comparison cycle, which improves the conversion speed by around 2 times.

The most obvious benefit of utilizing 2bit/cycle architecture is the reduction in number of interleaving channels. In high speed ADCs, the design challenge today lies in making interleaving work well. Another benefit of 2bit/cycle architecture is the reduction in sampling

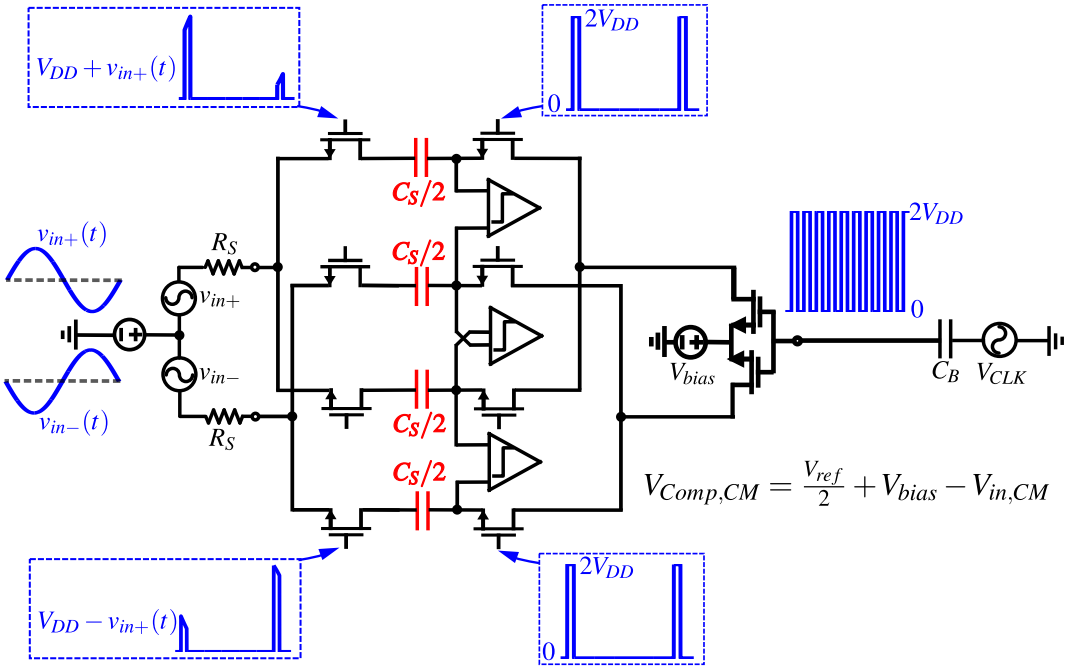
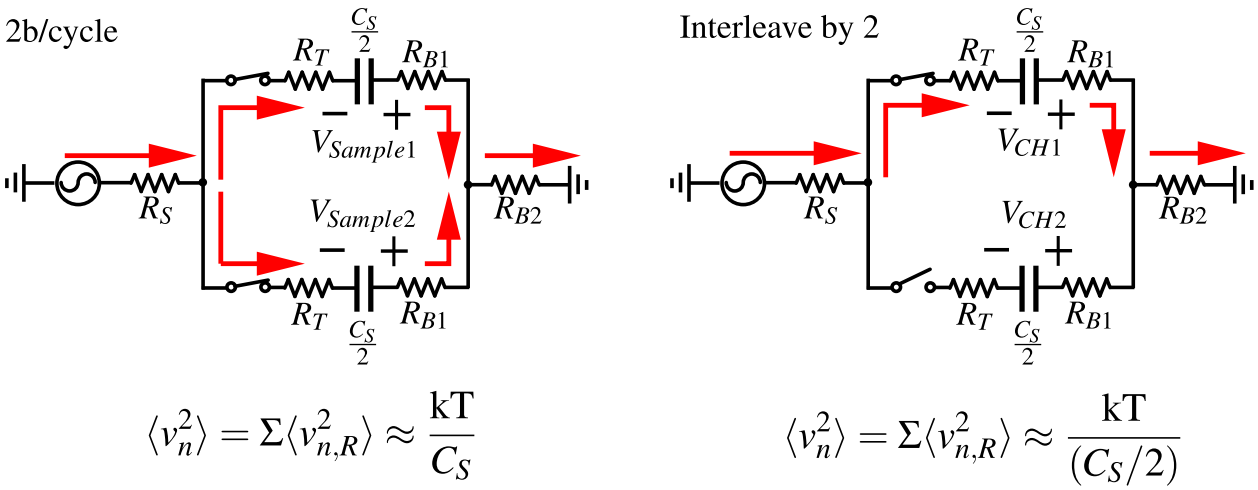


Figure 4.13: Differential 2bit/cycle SAR ADC using interpolation to remove one capacitive DAC.

thermal noise. Fig.4.14 compares the noise in the case of directly interleaving by 2 and the case of applying 2bit/cycle configuration. They have the same sampling capacitance C_S in CDAC and operate at the same sampling speed to the first order. In the case of interleaving by 2, only $C_S/2$ within one sub-channel samples the input signal and thus the thermal noise power is $\mathrm{kT}/(C_S/2) = 2\mathrm{kT}/C_S$. While in the case of 2bit/cycle configuration, both CDACs participate in tracking and the noise is reduced to kT/C. This reduction in thermal noise is valuable in applications where total capacitance is dominated by thermal noise. In these applications, beyond reducing number of interleaving channels, 2bit/cycle configuration also saves 50% consumption compared to directly interleaving 1bit/cycle SAR ADCs by reducing the total capacitor and its associated digital logic circuits, which can take up half of the total power consumption in high speed applications.

Fig.4.15 demonstrates plate voltages of a 2bit/cycle SAR ADC with 2 CDACs generating 3 pairs of differential voltages by interpolation. These plates are connected to the three comparators Fig.4.12. On the top are voltages of the four plates from two CDACs, which



$$\langle v_n^2 \rangle = \Sigma \langle v_{n,R}^2 \rangle \approx \frac{kT}{C_S} \qquad \langle v_n^2 \rangle = \Sigma \langle v_{n,R}^2 \rangle \approx \frac{kT}{(C_S/2)}$$

Figure 4.14: Comparison of kT/C noise between (a) 2bit/cycle configuration (b)and directly interleaving by 2.

converge to approximately the same voltage at the end of conversion phase (<1 LSB). In the middle are the three pairs of differential voltages with one pair generated by interpolation. They converge to the same voltage around zero (<1 LSB) at the end of conversion phase. At the bottom are the common-mode voltages applied to the comparators. Because of interpolation, one comparator experiences a changing common-mode voltage while the other two comparator have a constant common-mode voltage. This creates a unique challenge for comparator design as comparators demand well defined stable common-mode voltages to suppress offset and noise.

4.6.1 Capacitor DAC Design

The CDAC introduces 72LSB redundancy to deal with the changing common-mode voltage. Before the common-mode voltage approaches constant, the redundancy recovers comparator decision errors caused by the changing common-mode voltage. In the last LSB bits, all comparators have stable common-mode voltages. In total, the ADC resolves 10 bits in 6 cycles with 72LSB redundancy inserted.

Instead of purely scaling capacitance in the capacitor DAC, the reference voltage is also scaled to reduce the capacitance ratio between MSB capacitor and LSB capacitor. Targeting

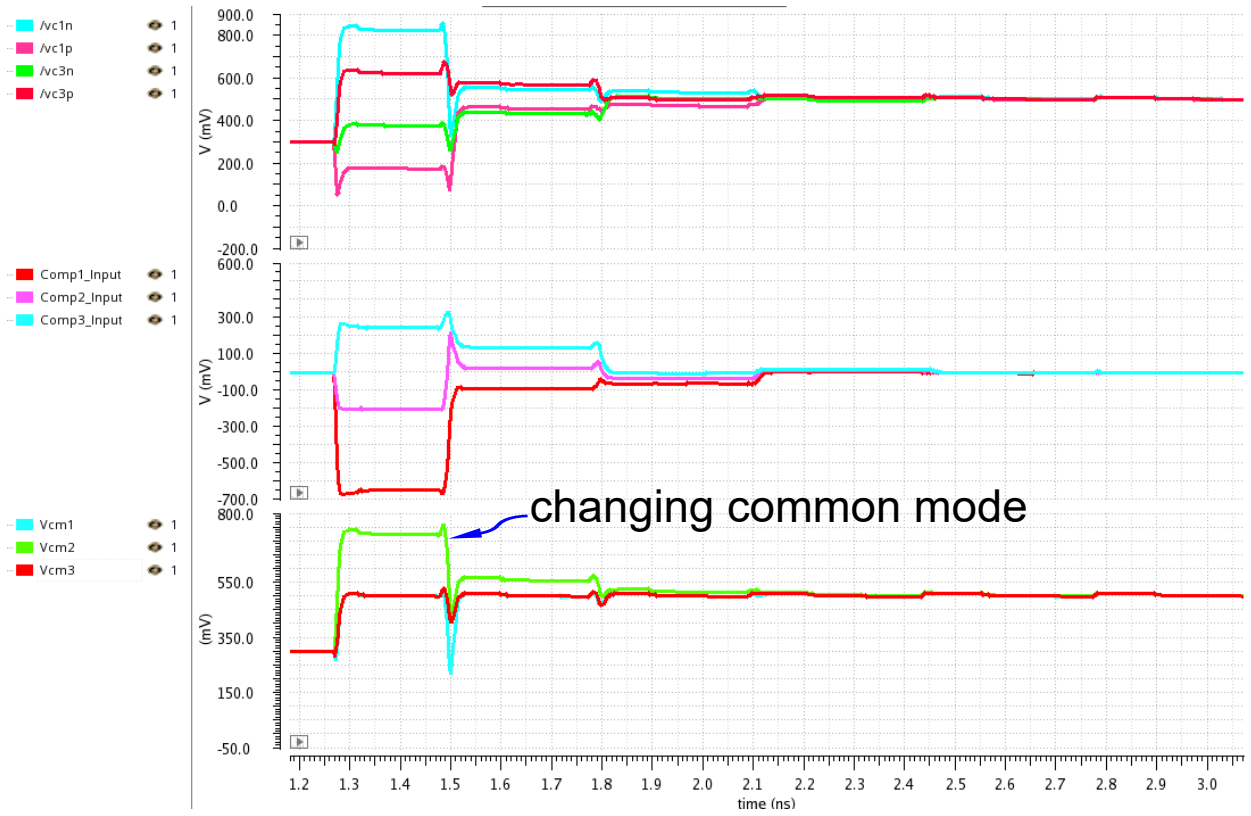


Figure 4.15: Transient plate voltages through conversion phase in a 2bit/cycle charge redistribution SAR ADC.

10bit physical resolution requires scaling LSB unit capacitor C_0 by $2^9 = 512$ times for MSB capacitor and a total capacitance of $2^{10} = 1024C_0$. Two dimensional common centroid capacitor layout is often essential to satisfy the matching requirement. In this design, besides V_{ref} and GND serving as positive and negative references, $\frac{V_{ref}}{4}$ and $\frac{V_{ref}}{16}$ are also utilized. This reduces MSB capacitor to $16C_0$ and enables compact one dimensional horizontally capacitor alignment as shown in Fig.4.16. For fast DAC settling, $\frac{V_{ref}}{2}$ is avoided as it needs complementary switches everywhere else. Only PMOS or NMOS switch are needed to drive each capacitor. One LSB is then $(C_0 \cdot \frac{V_{ref}}{16})/\Sigma C$. Total DAC resolution with redundancy is configured as below.

$$512 + 256 + 128 + 64 + \underbrace{64}_{\text{Redundancy}} + 32 + 16 + 8 + \underbrace{8}_{\text{Redundancy}} + 4 + 2 + 1(\text{LSB})$$

Unit capacitor is implemented with custom metal-to-metal(MOM) capacitor because

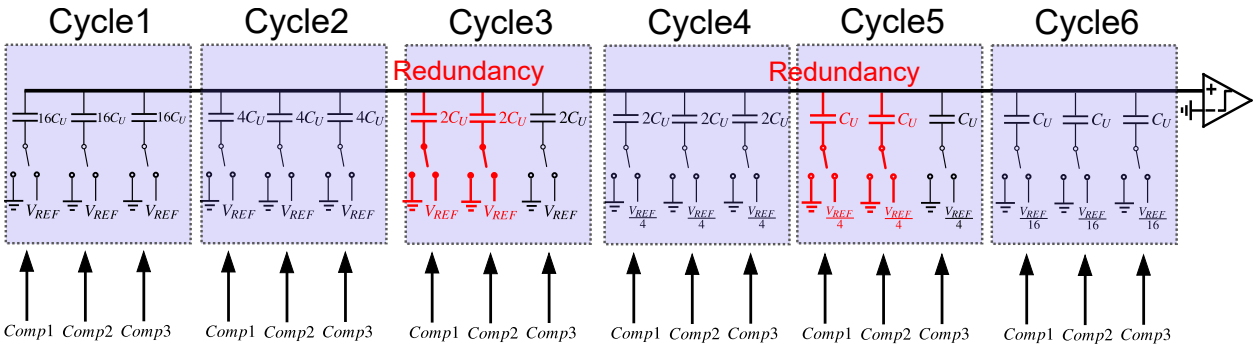


Figure 4.16: 10bit capacitor DAC with 72LSb redundancy and fractional reference.

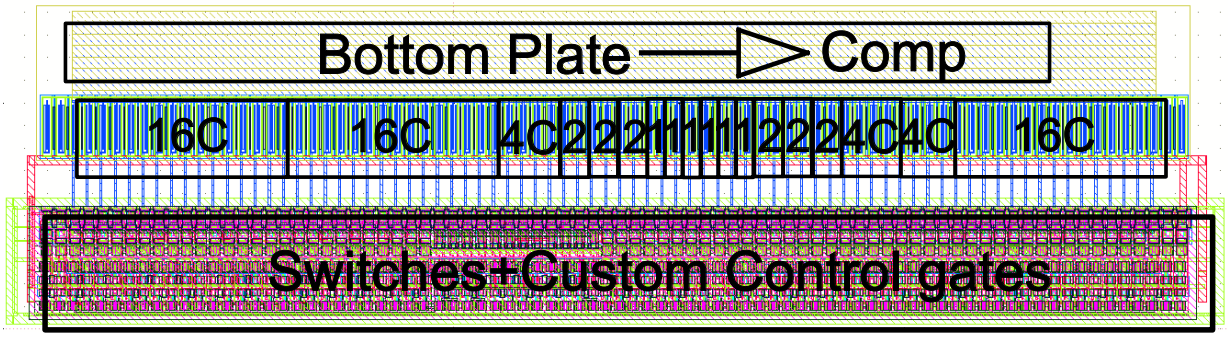


Figure 4.17: Layout of capacitor DAC.

MOM cap has better linearity and matching properties. Unit capacitance is chosen as $C_0 = 2$fF. MSB capacitor is $32C_0 = 64$fF. Each CDAC has a total capacitance of 156fF. With two capacitor DACs, the total capacitance presented to the T/H becomes 312fF. This leads to a thermal RMS noise of 0.12mV from track/hold, which does not contribute significantly to the overall noise.

4.6.2 Asynchronous SAR logic

In 2bit/cycle architecture, three comparators compare three thresholds and provide three bits of thermometer output code to the DAC. Fig.4.18 shows detailed diagram of 2bit/cycle SAR asynchronous control. The control logic first pre-sets the two DACs to create three pairs differential voltages for comparators. After all comparators generate full logic level outputs, a flag signal triggers the shift registers to provide sequential clock signals, which further act as clocks for SAR control logic circuits. When comparators are reset, their ready

signals are all '1's, the NOR gate outputs '0'. Only when comparators all finish comparisons and output '0's as ready signals, the NOR gate outputs a '1' to trigger the digital logic. Based on the outputs from the comparators, the control logic then decides the output of the current cycle and meanwhile pre-sets capacitor DAC for the next cycle.

As highlighted in Fig.4.18, correct operation of the asynchronous logic relies on alignment of delays of two paths: delay of self-triggered comparators t_{comp} and the delay from SAR control logic to capacitor DACs $t_{SR}+t_{SAR}+t_{DAC}$. The comparators should only be triggered after the capacitor DAC completely settles, otherwise comparators would compare incorrect plate voltages and output wrong decisions. Artificial delay stages with t_{delay} must be inserted to align the comparator triggering time with the delay from digital logic to capacitor DAC, including delay of the shift register t_{SR}, delay of the SAR control logic circuits t_{SAR}, which mainly consists of D flip-flops and RC settling time of the capacitor DAC t_{DAC}. $t_{delay} > t_{SR}+t_{SAR}+t_{DAC}$ must be satisfied. The delay line is implemented with fixed delay for worst case CDAC settling time.

The conversion speed of the SAR ADC is ultimately determined by $t_{comp}+t_{SR}+t_{SAR}+t_{DAC}$. t_{comp} is the time for comparators to resolve the inputs and is a function of their input amplitudes. Overall, one cycle of conversion is about 150ps, leading to about 900ps total conversion time for six cycles. After all cycles are finished, outputs are stored by D flip-flops in the data synchronization circuits for further interleaving. Only standard cells were used in this design to save labor. Implementing all controls with custom logic can further speed up the conversion by around 30%.

4.6.3 Calibrated low-offset low-noise comparators

StrongARM comparator in Fig.4.19 is chosen in this design for its compactness, low-power operation and inherent voltage amplification that suppresses offset and noise [22]. Binary scaled calibration capacitor arrays are inserted into both C_C and C_L. C_C has stronger ability of correcting offset and does not impact regeneration constant on first order while C_L is used for fine offset tuning. These calibration capacitors are implemented as MOS capacitors

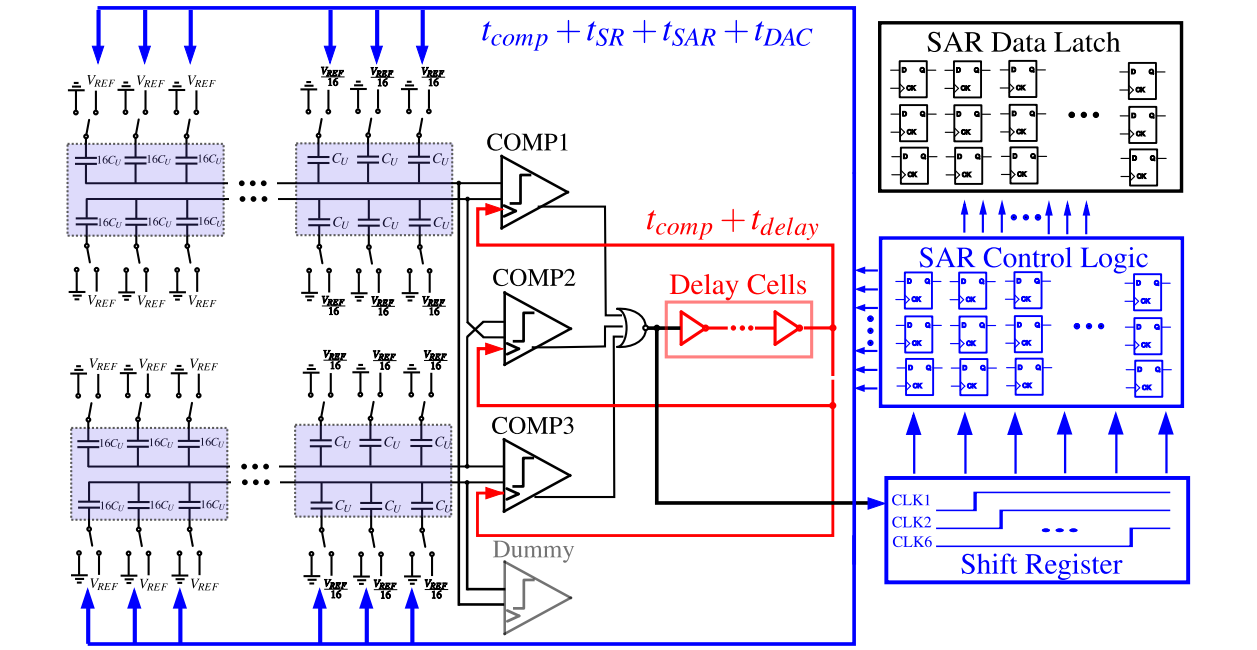


Figure 4.18: 2bit/cycle SAR with asynchronous SAR control logic.

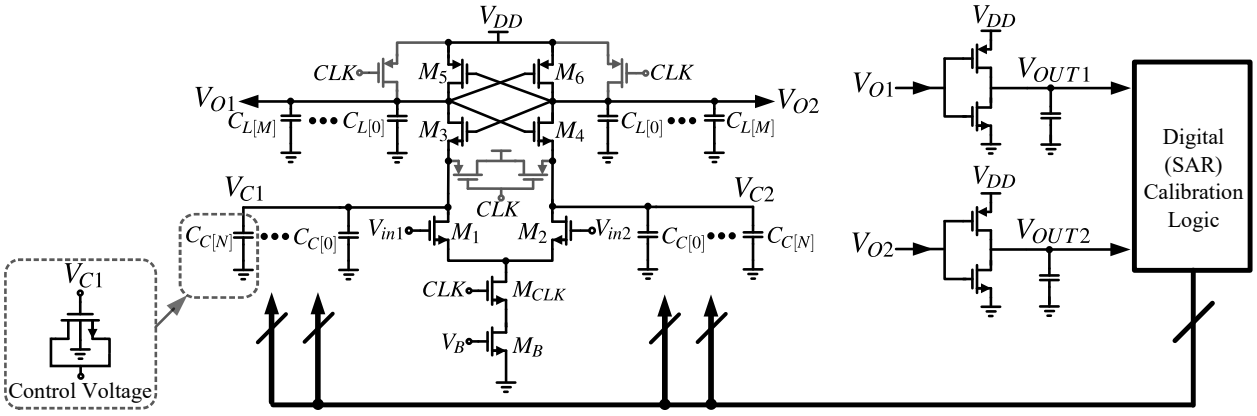


Figure 4.19: Schematic of the applied strongARM comparator with offset calibration.

with drain-source controlled by binary output from digital logic circuits. The digital control logic for offset calibration is essentially a SAR logic. After shorting the input with a fixed common-mode voltage, the control logic programmes the weight ratio between C_{C1}, C_{C2} and C_{L1}, C_{L2} depending on the comparator output.

The maximum offset that the calibration scheme can cover is limited by the ON/OFF ratio of the MOS cap, $\sim$2 in this technology. Maximum capacitor added at C_C and C_L, besides wirings, are 20fF and 4fF respectively. C_C consist of 6bit binary scaled capacitors

with MSB capacitor of 10fF while C_L consists of 5bit binary scaled capacitors with MSB capacitor of 1fF. Majority of calibration capacitors are added at C_C to avoid degrading regeneration constant. Comparators are driven by input common-mode of 550mV. Thus, the maximum offset that can be calibrated is

$$V_{OS,max} = \underbrace{\frac{2\Delta C_C + \Delta C_L}{C_L + 2C_C}}_{\text{ratio}} \cdot \underbrace{\frac{V_{OV1,2}}{2}}_{\text{bias}} \sim \mathbf{32mV} \tag{4.15}$$

where ΔC_C = 10fF, ΔC_L = 4fF, C_C = 26fF including intrinsic capacitors, C_L = 10fF including intrinsic and loading capacitors, $V_{OV,1,2} = 550 - 300 - 50 = 200$mV. From Monte-Carlo simulation, the comparator without calibration has a standard deviation of σ =5mV in offset distribution. The calibration capacitors can cover 5σ spread of offsets.

The highest resolution, or the finest tuning step of the offset calibration scheme, relies on the minimum capacitor of the technology. This is the residual offset after calibration is done. The smallest cap used is a MOS cap of W/L =80nm/30nm size with equivalent capacitance of 0.125fF at C_L. This leads to a resolution of

$$V_{OS,min} = \underbrace{\frac{C_{L,min}}{C_L + 2C_C}}_{\text{MOS cap}} \cdot \underbrace{\frac{V_{OV1,2}}{2}}_{\text{bias}} \sim \mathbf{0.25mV} \tag{4.16}$$

Fig.4.20 plots the offset distribution before and after calibration from 300-run Monte-Carlo simulation. Before calibration, the comparator has a standard deviation of 5mV for offset with maximum offset between [-23mV, 20mV]. After calibration, the offset is bounded below 0.4mV. It is a bit larger than the predicted value above because calibration capacitors also have mismatches. However, with input full swing of 1.6V and 1LSB of 1.6mV, the residual offset would not impact the overall SNDR.

While offsets can be calibrated by unbalancing capacitor loadings, noise can be only reduced by increasing the absolute amount of loading capacitors. The added calibration capacitors serve not only as a method of correcting offset, but is also inevitable for noise reduction. They are therefore necessary for two reasons. Input referred noise of the comparator

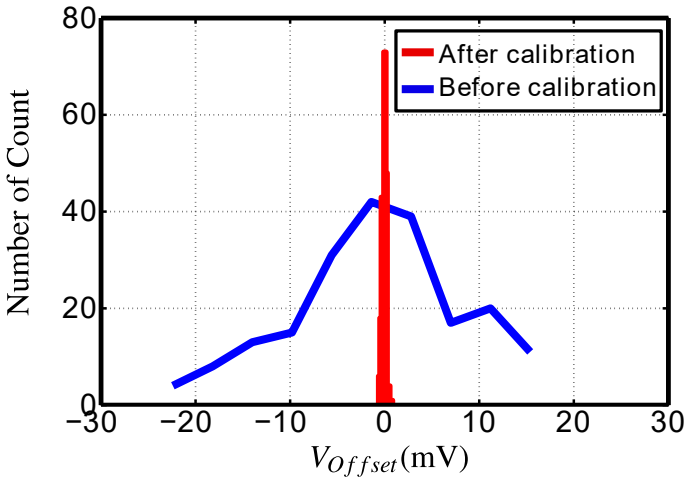


Figure 4.20: Monte-carlo simulation of offset distribution of the comparator before/after calibration.

with calibration capacitors is already derived (2.50) in Chapter 2.

$$\begin{aligned}\langle v_{n,in}^2 \rangle &= 2\mathrm{kT}\gamma \cdot \frac{V_{OV1,2}}{V_t} \cdot \frac{1}{C_L + 2C_C} \\ &= \underbrace{\frac{\mathrm{kT}}{C_L + 2C_C}}_{\mathrm{kT/C\ noise}} \cdot \underbrace{\frac{V_{OV1,2}}{V_t}}_{\mathrm{bias}} \cdot \underbrace{2\gamma}_{\mathrm{constant}}\end{aligned} \tag{4.17}$$

Shown in Fig.4.21 is the simulated input referred RMS noise versus input common-mode voltage. A set of fixed differential voltages are applied to the comparator in transient noise simulations. Without noise, the comparator would output purely '1's and '0's. The output distribution with added transient noise is then fitted to a Gaussian distribution whose sigma is also sigma the input referred noise voltage. This calculation assumes the input referred noise has a Gaussian distribution. At $V_{in,CM} = 550\mathrm{mV}$, the comparator achieves the optimal trade-off between noise and speed. Reducing input common-mode results in slower operation while further raising the input common-mode voltage degrades noise as input pair would be pushed into triode region. This corresponds to an input referred RMS noise of 0.3mV, which falls well below 1LSB.

Fig.4.22 shows the calibrated comparators. Besides the core comparator and calibration logic, buffers consisting inverters are added and placed close to the comparator to drive the

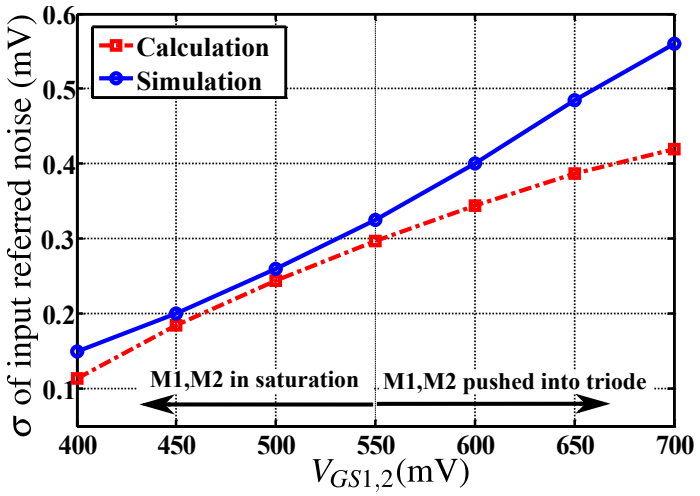


Figure 4.21: Simulated input referred RMS noise versus input common-mode voltage.

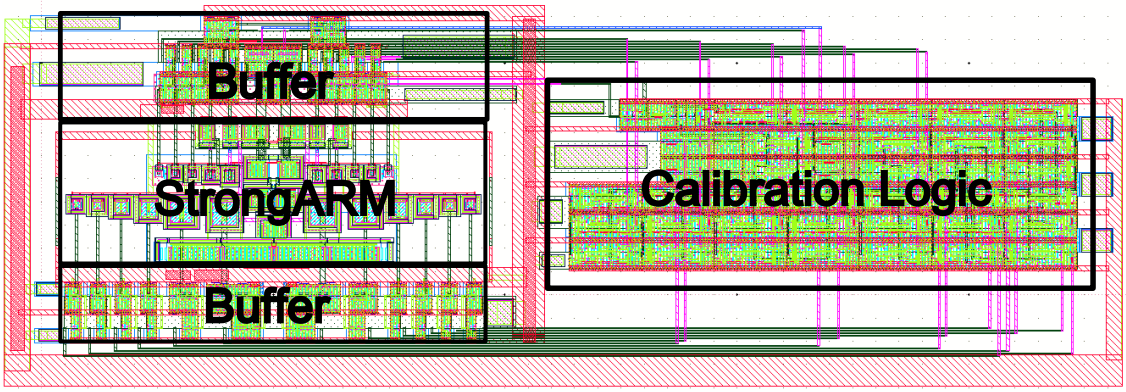


Figure 4.22: Layout of the comparator with calibration circuits.

drain/source of the calibration MOS caps. They are necessary to hold the control voltages steady so that switching from the logic would not interfere with the comparator.

Fig.4.23 shows the layout of a single sub-channel SAR ADC. It measures 120μm×220μm. Its area can be further reduced with better layout. This design was finished in a short time period, thus its layout is not fully optimized.

4.7 Interleaving Sub-channel ADCs

Fig.4.24 shows the conceptual diagram of interleaving eight sub-channel ADCs with a single master switch. Position of sub-channel 1-8 in the diagram also denotes the actual layout

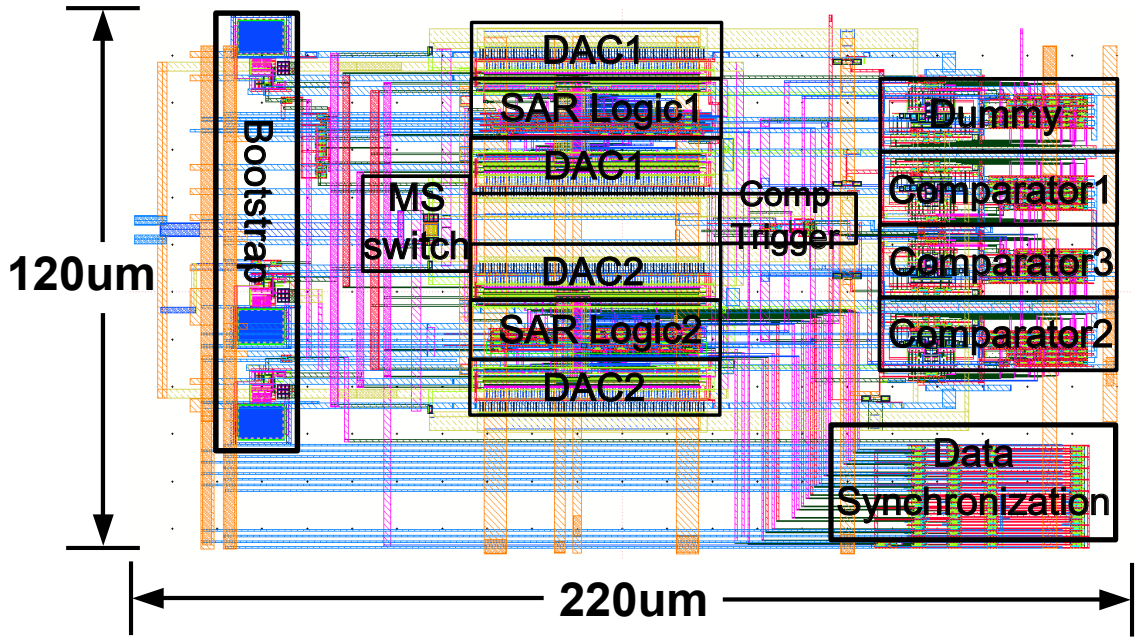


Figure 4.23: Layout of the sub-channel 500MS/s 10bit ADC.

alignment. They are positioned to provide symmetric routing for each sub-channel ADC. This implementation removes all timing skews as the sampling instants will be solely determined by the master switch. However, the long routing wires connecting the master switch to sub-channels severely degrades tracking bandwidth as shown in more detailed circuits in Fig.4.25. Using a single master switch with long routing wires worsens the performance of the interleaved ADC in several aspects

- High tracking bandwidth is always preferred as it is essential to suppress bandwidth mismatch. The impact of bandwidth mismatch will be discussed in details in the following sections.

- Junction caps from other off-channels also adds capacitive loading to the active channel. This not only decreases the tracking bandwidth, these junction caps are non-linear capacitors themselves and are a source of distortions.

- Signal swings across the sampling switches become higher due to wiring RCs, which further leads to higher distortions.

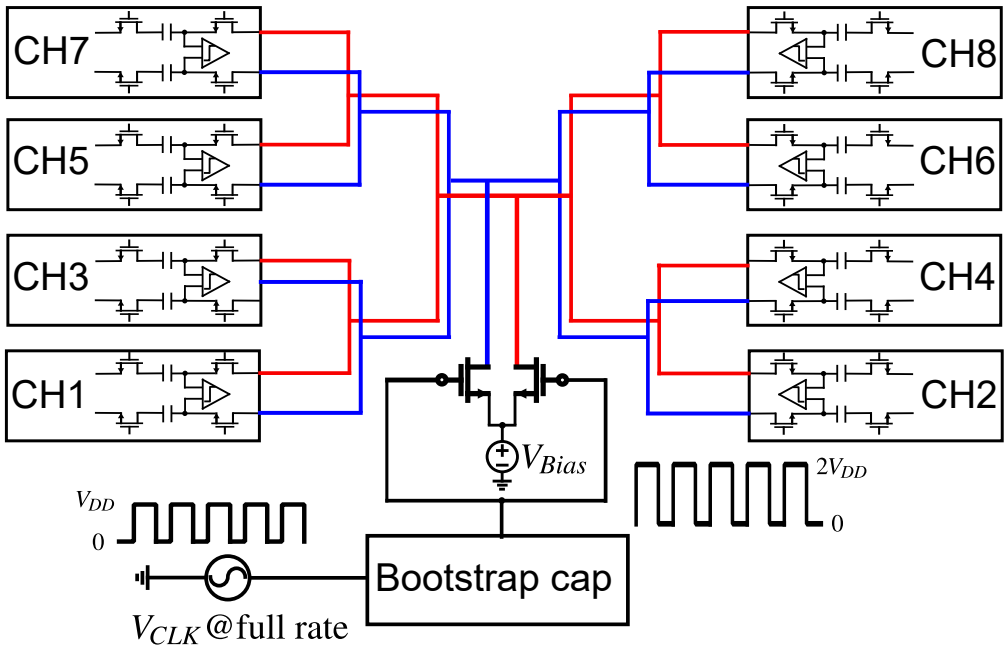


Figure 4.24: Master sampling switch implemented as a single switch.

Connecting the sub-channel ADC to the master switch would require interconnect wires with length of 800μm. This translates to an equivalent resistance of 50Ω of the wire alone, let alone the distributed parasitic capacitors along the interconnect, which is hard to predict because of other adjacent wires. As discussed in previous sections, the total resistance including source resistor and all sampling switches are already 100Ω, the additional 50Ω from wirings would degrade the tracking bandwidth to around 3GHz. These negative impacts from wiring offsets the benefit of master/slave sampling scheme and demands a better solution of interleaving these sub-channel ADCs.

Distributing the master sampling switch into each channel solves the problem by removing the wiring parasitics from the signal path as shown in Fig.4.26. This arrangement also isolates the active channel from junction capacitors from other off channels. Instead of loading the sampling switches, now the wiring loads only the boot-strapped circuit, which drives the gates of the distributed master switches and is outside the critical signal path Fig.4.27. By increasing driving strength of the boot-strapped circuit, sharp rise/fall edges of the boot-strapped clock are still maintained despite the impact from wiring parasitics. This is relatively easy to achieve as equivalent gate capacitance of the distributed sampling

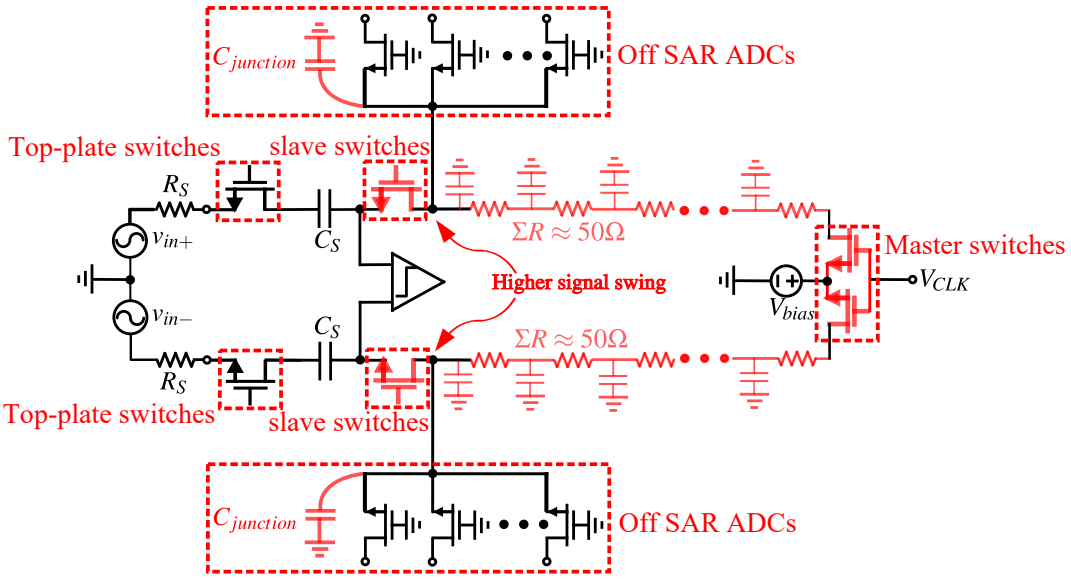


Figure 4.25: RC parasitics from long routing wires severely limit tracking bandwidth.

switches is below 15fF.

The distributed master switches, however, results in another problem. Although they share the same master sampling clock, threshold mismatch among these switches will result in timing skew. But we will show in the following section that this timing skew is small enough so it does not degrade SNDR of the interleaved ADC. Layout is arranged to match wirings of all sub-channels to minimize potential bandwidth mismatches.

All interleaving sub-channel ADCs share the same global reference ladder to avoid gain mismatch caused by difference references as shown in Fig.4.29. The reference ladder is implemented as a piece of metal with total equivalent resistance of 30Ω. This ensures that the reference ladder will not limit the settling speed of the capacitor DAC. The reference ladder has on-chip decoupling MOM capacitors of 500pF which aims to minimize the voltage ripple caused by switching currents from the capacitor DAC.

Layout of the core circuitry is shown in Fig.4.30 including 8 interleaving sub-channel ADCs and a decimator. The decimator combines the data from 8 sub-channels downsamples them to 40MHz. Without the decimator, full rate 2GHz output data would be almost impossible for typical PCB board to support.

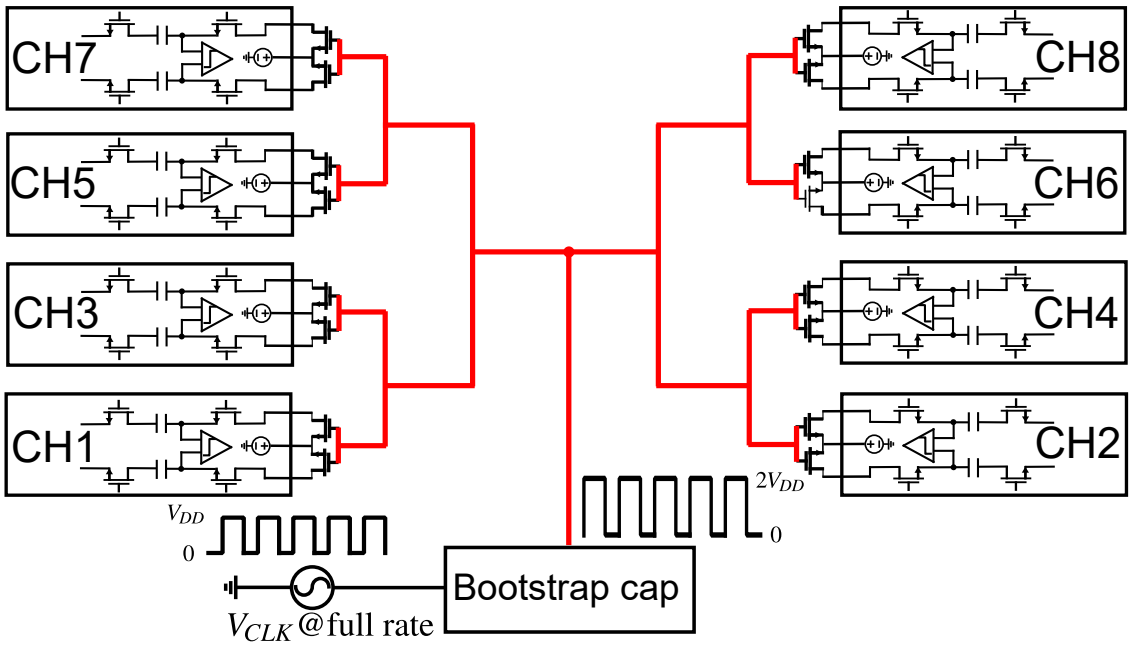


Figure 4.26: Master sampling switch is distributed into each sub-channel ADC.

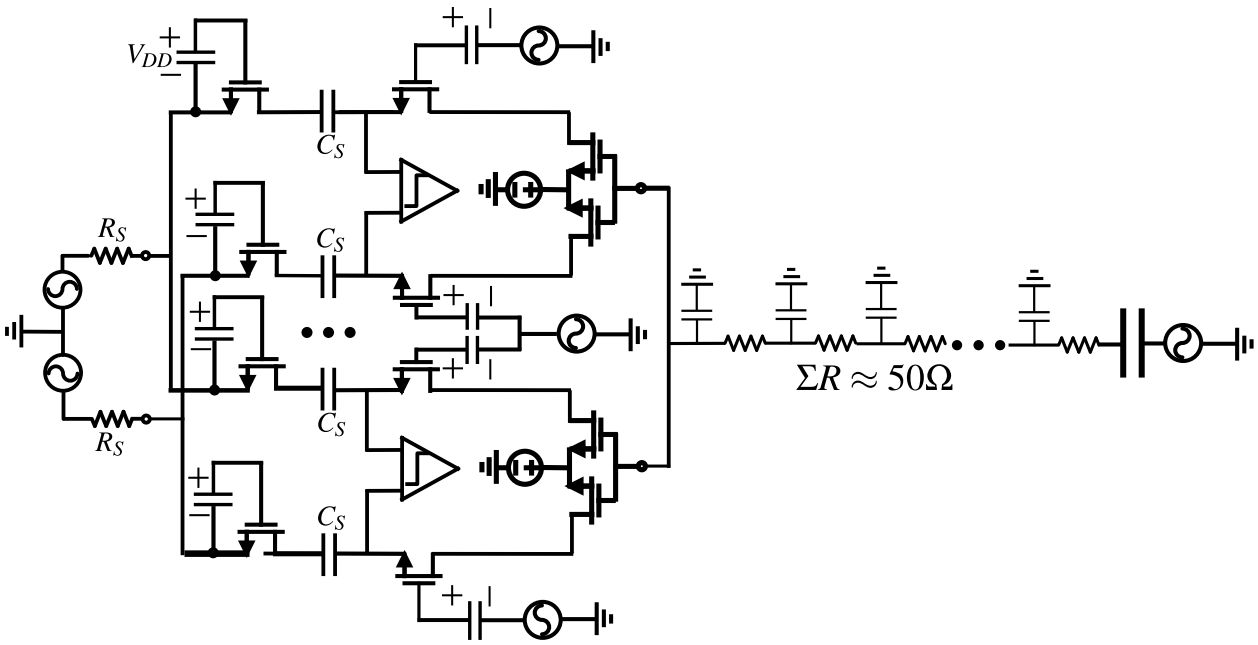


Figure 4.27: RC parasitics from long routing wires removed from the critical signal path.

Layout of the full chip is shown in Fig.4.31. The chip measures 1.5mm×1.5mm area with core circuitry occupying 730μm×820μm. The rest of the area is filled with MOM decoupling capacitors, which also helps fulfill the density requirement.

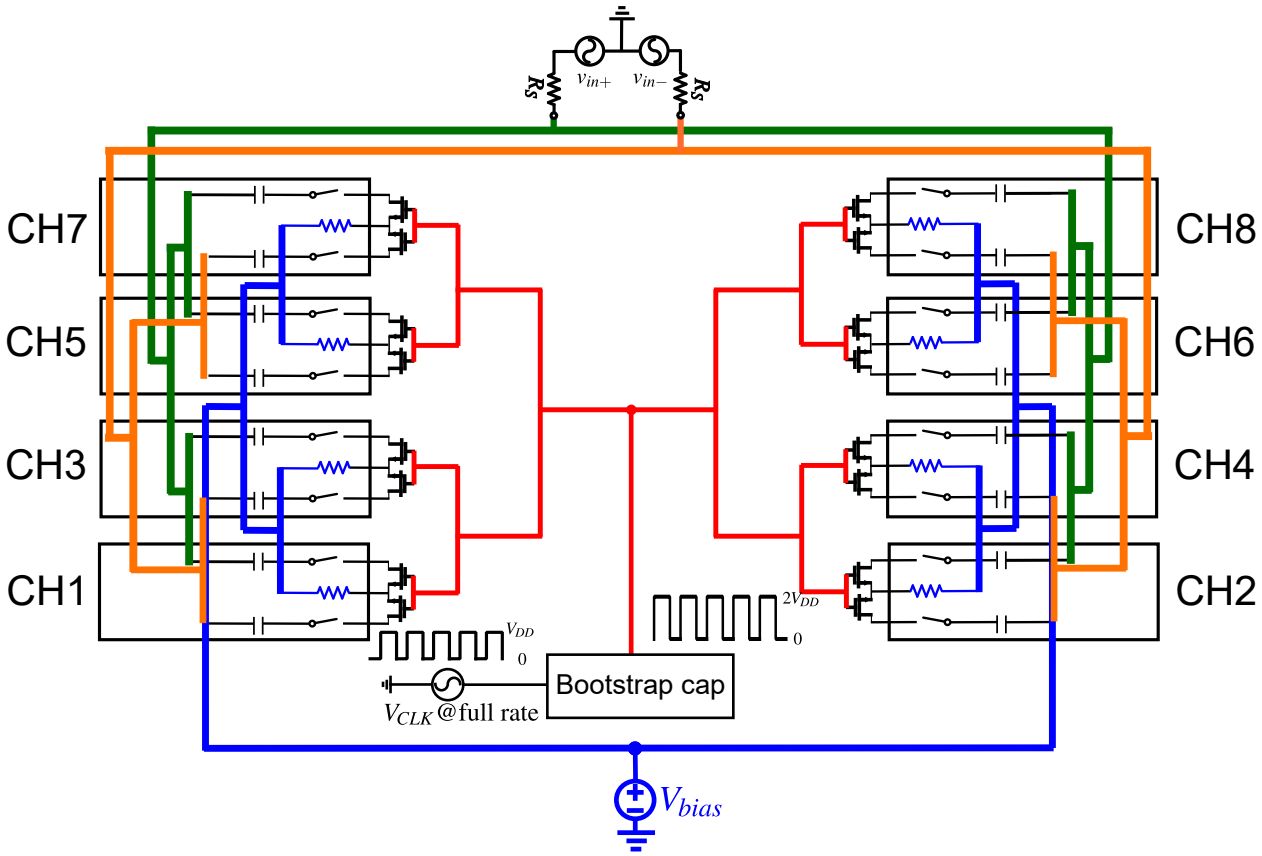


Figure 4.28: Diagram showing symmetric wiring of the critical signal path.

4.8 Sources of additional degradation

4.8.1 Timing skew caused by V_T mismatch among distributed master switches

Timing skew causes the input signal to be non-uniformly sampled by sub-channels. In an N-channel interleaved ADC, the SNDR set by timing skew [122] is

$$\mathrm{SNR}_{\Delta t} = \frac{N}{N-1} \cdot \frac{1}{|R''(0)|\langle \Delta t^2 \rangle} \tag{4.18}$$

where $|R''(0)|$ is the second derivative of the autocorrelation function $R(0)$ of the input signal. For the single tone sinusoidal test where the input is pure sinewave $x(t) = \sin(2\pi f_{in} t)$, $R''(0) = -(2\pi f_{in})^2$, thus

$$\mathrm{SNR}_{\Delta t} = \frac{N}{N-1} \cdot \frac{1}{(2\pi f_{in})^2 \langle \Delta t^2 \rangle} \tag{4.19}$$

This provides a quantitative estimate of the impact of timing skews.

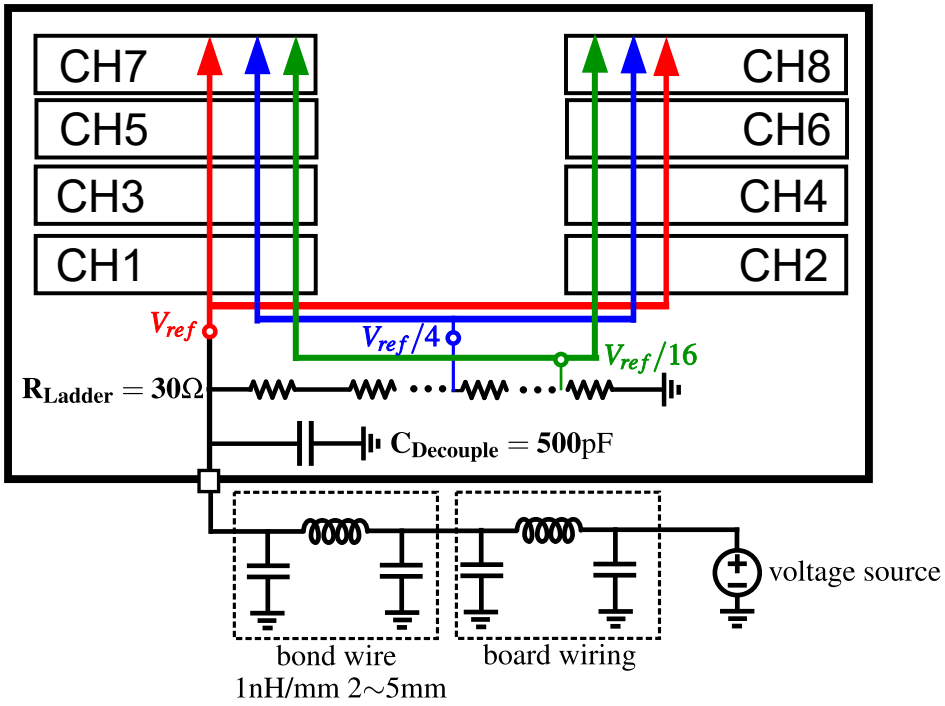


Figure 4.29: Global reference ladder for all interleaving sub-channel ADCs.

Standard deviation of the threshold mismatch among the distributed master switches is calculated as

$$\sigma_{\Delta V_t} = \frac{A_{V_t}}{\sqrt{W \times L}} = \frac{2\text{mV} \cdot \mu\text{m}}{\sqrt{18\mu\text{m} \times 0.03\mu\text{m}}} = 2.7\text{mV} \tag{4.20}$$

where A_{V_t} is taken from the PDK, 1.8um/30nm is the size of the distributed switches. A sampling clock slope of 1.5V/25ps translates this to a timing skew standard deviation of

$$\sigma_{\Delta t} = \frac{\Delta V_t}{\mathrm{d}V/\mathrm{d}t} = \frac{2.7\text{mV}}{1.5\text{V}/25\text{ps}} = 45\text{fs} \tag{4.21}$$

The worst case occurs when input frequency approaches Nyquist frequency $f_{in} = 2\text{GHz}$,

$$\begin{aligned} \text{SNR}_{\Delta t} &= \frac{N}{N-1} \cdot \frac{1}{(2\pi f_{in})^2 \langle \Delta t^2 \rangle} \\ &= \frac{8}{8-1} \cdot \frac{1}{(2\pi \times 2\text{GHz} \times 45\text{fs})^2} = 65\text{dB} \end{aligned} \tag{4.22}$$

4.8.2 Bandwidth mismatch during tracking phase

Bandwidth mismatch arises from transistor mismatches among sampling switches, mainly β mismatch that causes mismatches among on-resistance of the sampling switches. It results

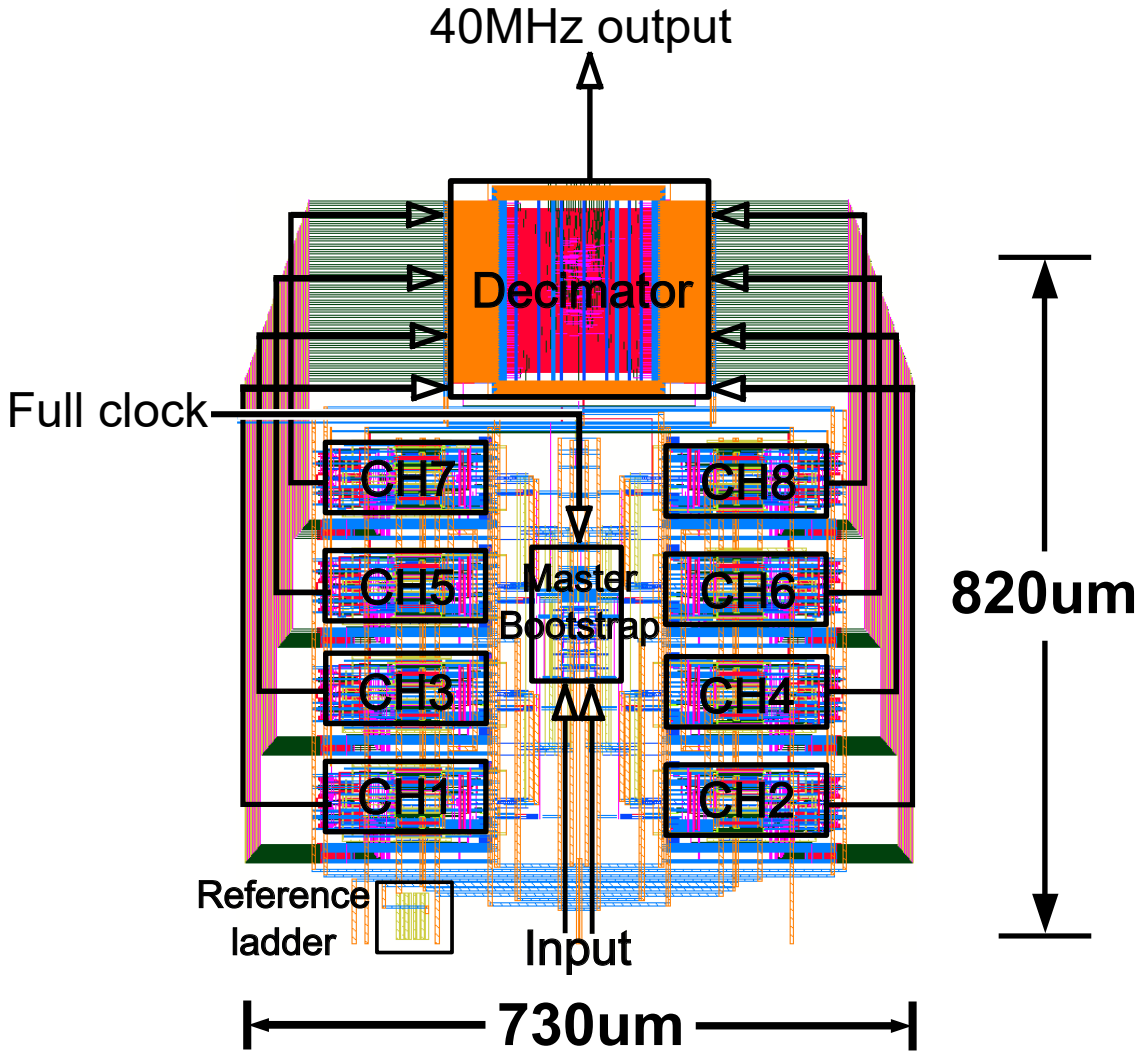


Figure 4.30: Layout of the core circuitry, including 8 interleaving sub-channel ADC and on-chip decimator.

in spurs that increase with the input (4.6). The bottom slave switch with the smallest size of the three switches, labeled with dashed lines in Fig.4.32, dominates bandwidth mismatch. Using $A_\beta = 0.4\% \cdot \mu\text{m}$ from the PDK, β mismatch can be calculated

$$\frac{\Delta\beta}{\beta} = \frac{A_\beta}{\sqrt{W \times L}} = \frac{0.4\% \cdot \mu\text{m}}{\sqrt{8\mu\text{m} \times 0.03\mu\text{m}}} = 0.8\% \tag{4.23}$$

The mismatch in on-resistance is then calculated

$$\frac{\Delta R_{ON}}{\Sigma R} = \frac{\Delta\beta}{\beta} \cdot \frac{R_{ON}}{\Sigma R} = \frac{20\Omega \times 0.8\%}{100\Omega} = 0.16\% \tag{4.24}$$

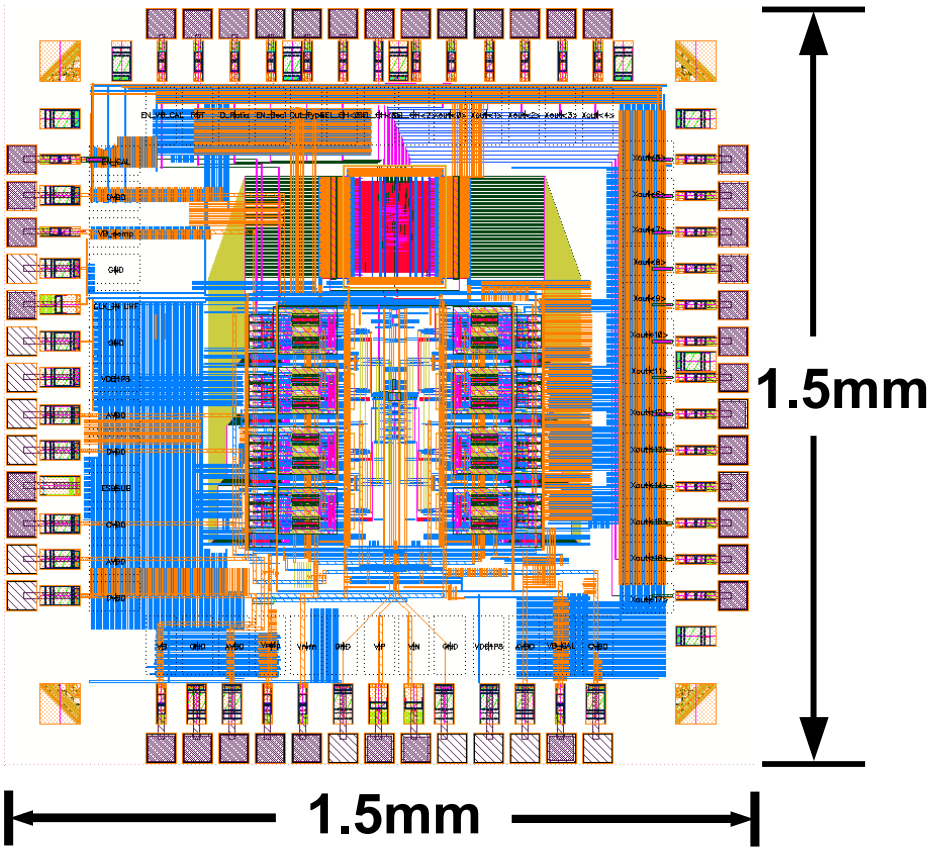


Figure 4.31: Layout of the full chip.

The equivalent timing skew due to bandwidth mismatch is then found by

$$\Delta t_{BW} = 0.16\% \times \frac{1}{2 \times \pi \times 5\text{GHz}} = 50\text{fs} \tag{4.25}$$

Again, assuming that sampling capacitors match perfectly, the worst case occurs when input frequency approaches Nyquist frequency $f_{in} = 2\text{GHz}$,

$$\begin{aligned} \text{SNR}_{\Delta t} &= \frac{N}{N-1} \cdot \frac{1}{(2\pi f_{in})^2 \langle \Delta t_{BW}^2 \rangle} \\ &= \frac{8}{8-1} \cdot \frac{1}{(2\pi \times 2\text{GHz} \times 50\text{fs})^2} = 64\text{dB} \end{aligned} \tag{4.26}$$

The simulated bandwidth mismatch is 0.19%, which matches well with calculation.

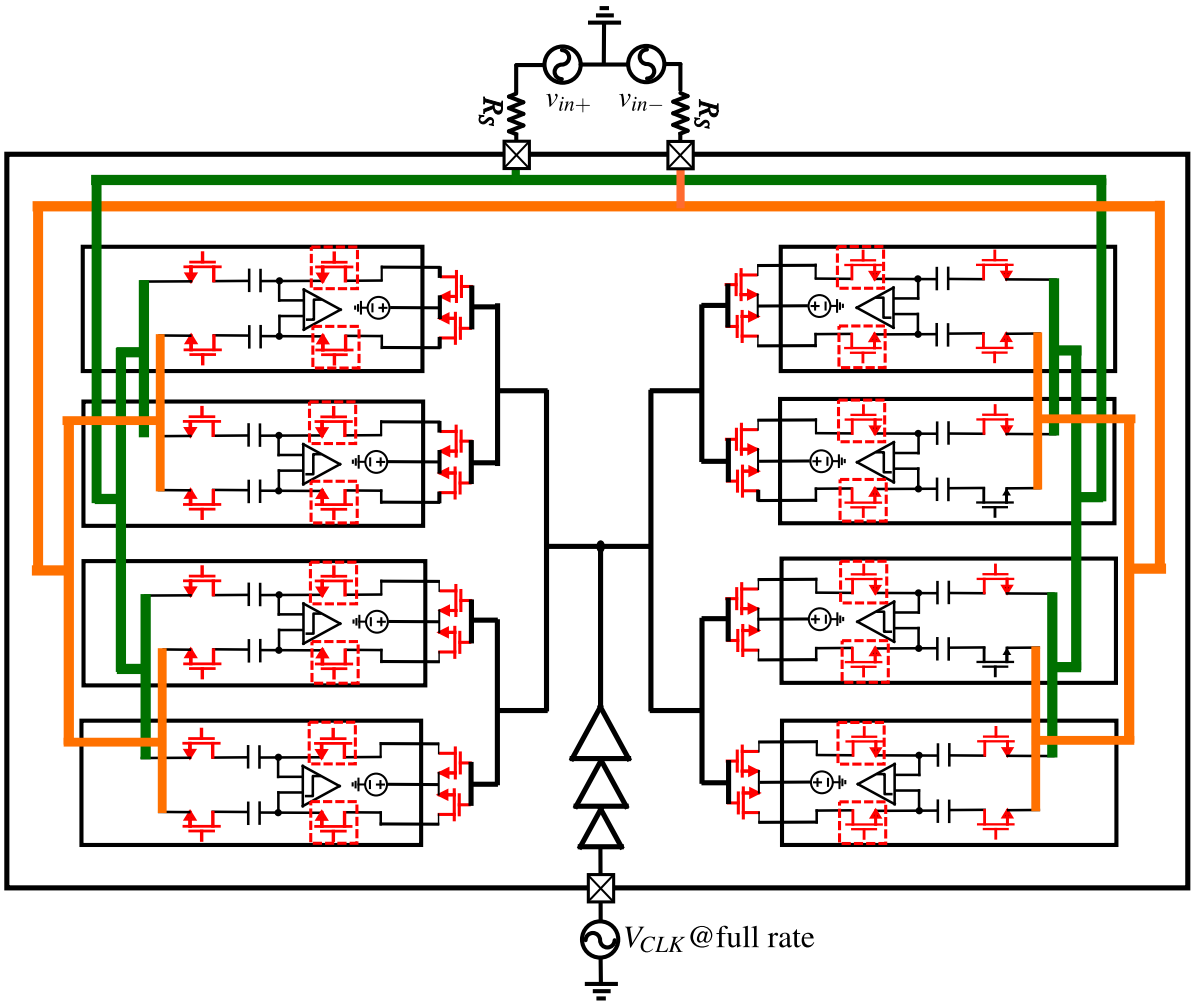


Figure 4.32: Bandwidth due to transistor mismatches.

4.9 Measurement and discussion

The supply voltage was reduced from designed 1.1V to 0.9V during measurement because of over-voltage limit requirement. The cause of this reduction will be discussed in later sections. The maximum sampling frequency is measured as 2.5GHz, above which the conversion cycles would not complete. Fig.4.33 shows the measured spectrum with low frequency input of 42MHz and sampling frequency of 2.5GHz. Interleaving spurs at $F_s/8 \pm f_{in}$, $2F_s/8 \pm f_{in}$ ... suggest gain mismatches among sub-channels while interleaving spurs at $F_s/8$, $2F_s/8$... suggest input referred offsets after calibration. SNDR at DC is 52.0dB and ENOB is 8.35bit. Fig.4.34 shows the measured spectrum with Nyquist input frequency of 1.2GHz and sampling

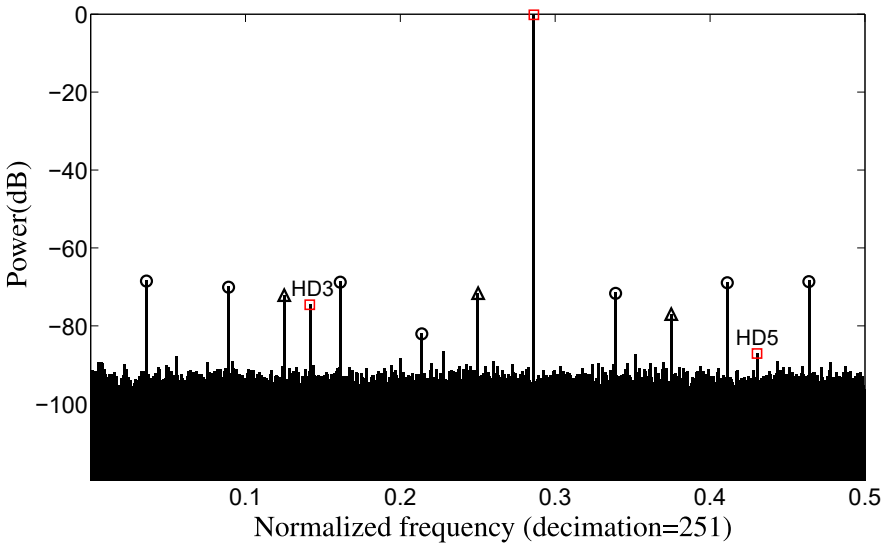


Figure 4.33: Measured FFT spectrum with $F_S = 2.5$GHz and F_{in}=42MHz.

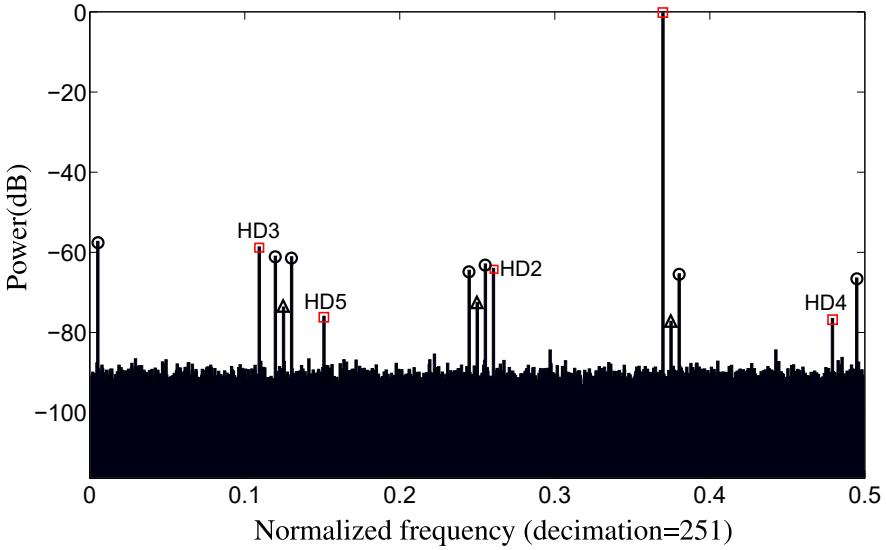


Figure 4.34: Measured FFT spectrum with $F_S = 2.5$GHz and F_{in}=1.23GHz.

frequency of 2.5GHz. Interleaving spurs at $F_s/8 \pm f_{in}$, $2F_s/8 \pm f_{in}$ ... have grown which are caused by timing skew/bandwidth mismatch. HD3 of 59dBc is the dominant harmonic distortion. SNDR at Nyquist input frequency degrades to 46.4dB and ENOB is 7.4bits.

Fig.4.35 is the measured INL/DNL of the overall interleaved ADC and Fig.4.36 plots INL/DNL of each sub-channel ADC, both measured with input frequency of 20MHz and sampling frequency of 630MHz. The peak DNL of the overall ADC is (-0.39LSB, 0.23LSB)

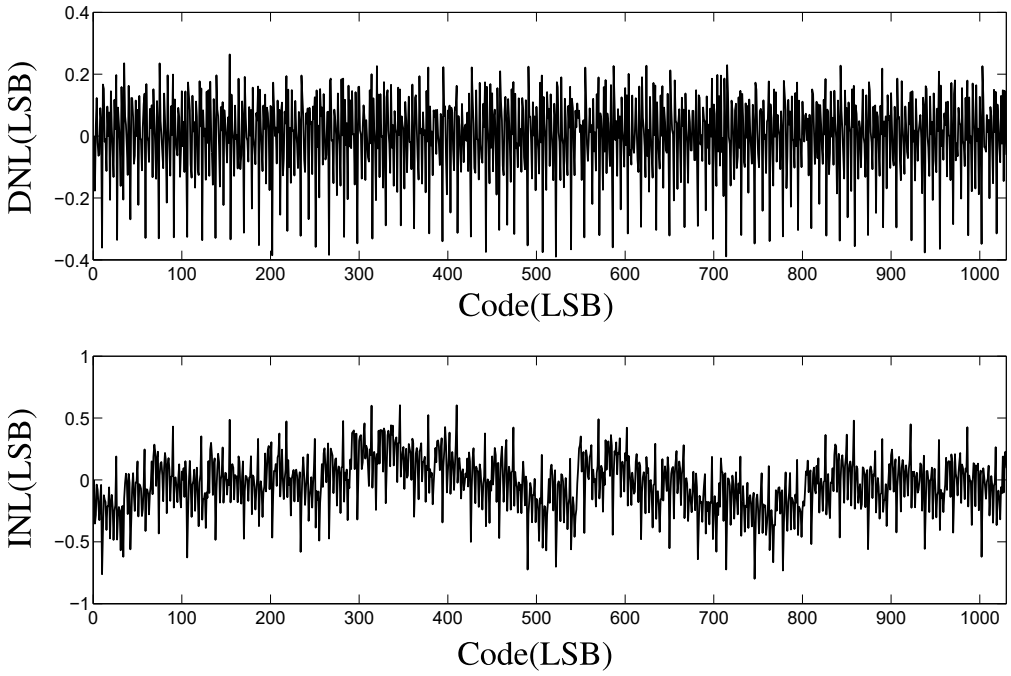


Figure 4.35: Measured DNL and INL with $F_S = 600$MHz and F_{in}=20MHz.

and the peak INL of the overall ADC is (-0.81LSB, 0.52LSB). INL/DNL of the overall interleaved ADC are significantly smaller because of averaging effect in the code density test where the same input voltage is applied to all sub-channels for multiple times. The measured INL/DNL is larger than expected and its reason will be explained in later sections.

Fig.4.37 plots SNDR versus the input frequency when the sampling frequency is fixed at 630MHz. Fig.4.38 plots SDNR, SFDR, HD3 versus input frequency respectively.

Fig.4.39 shows the measured power consumption versus input frequency. The dominant source of static power consumption is the reference ladder which is used to create fractional references. Rest of the circuits only consume dynamic power that scales with frequency. At the maximum sampling frequency of 2.5GHz, the reference ladder consumes 50% of the total power while clock buffers consume the second largest of 16mW, digital is the third consuming 10.4mW and analog power, mainly comparators, measures 4.8mW as the lowest power consuming block.

Table.4.1 compares the performance of the chip with state-of-arts. While SNDR and ENOB are comparable to most of the published works, power consumption is noticeably

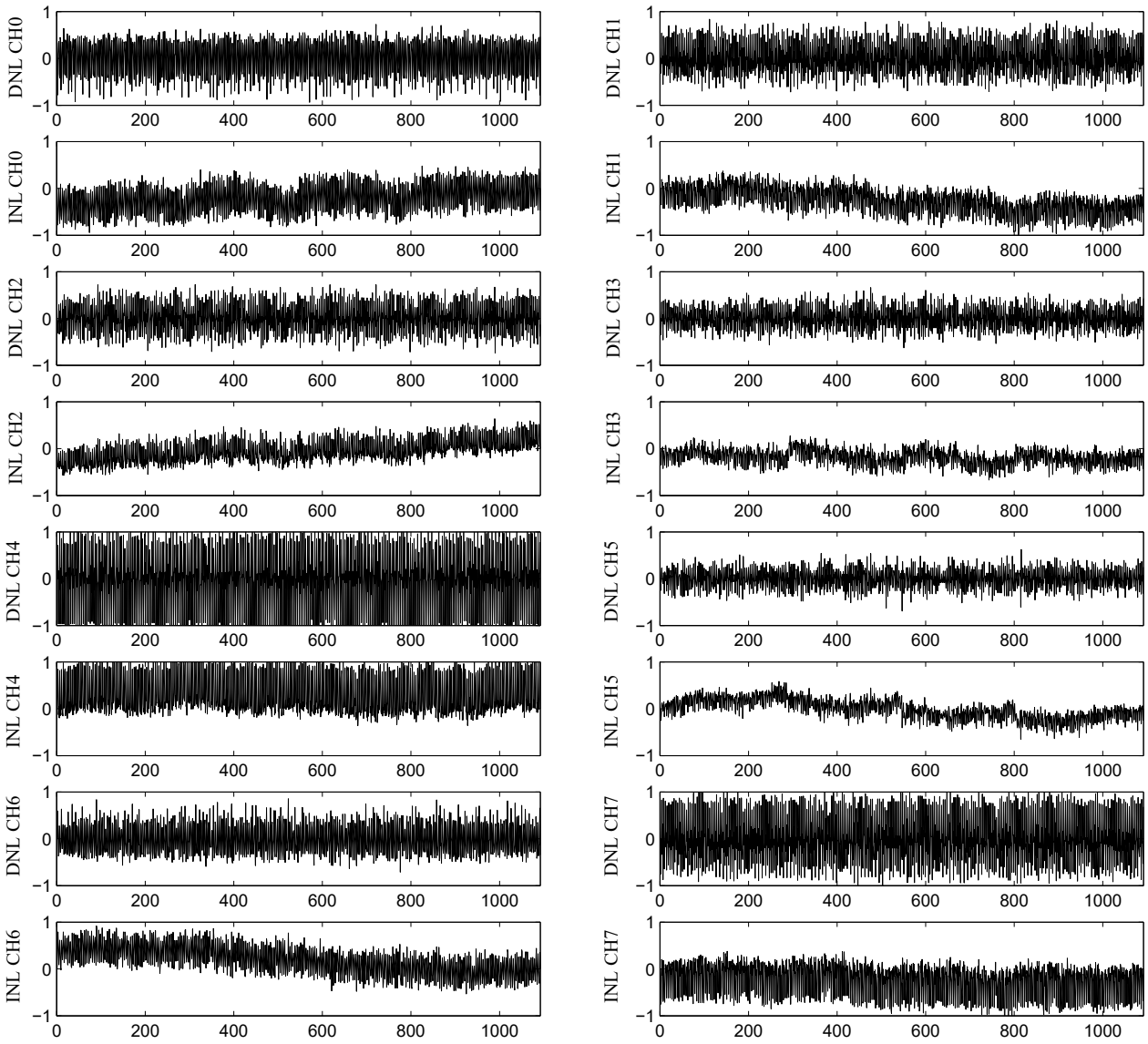


Figure 4.36: Measured DNL and INL with $F_S = 600$MHz and F_{in}=20MHz.

higher, which ultimately limits the figure-of-merit. As mentioned above, the dominant power consumption is the reference ladder. The current design is over conservative in the ladder power and later section will show that over 80% power reduction in this block will not limit the performance. Since digital circuits are all implemented with standard cells, further power reduction and speed improvement are both achievable by replacing standard cells with custom logic circuits in the SAR logic critical path. The clock buffers can further benefit from more optimized layout design that has shorter routing distance and smaller parasitics, but the power consumption of the clock buffers cannot be significantly reduced because it is

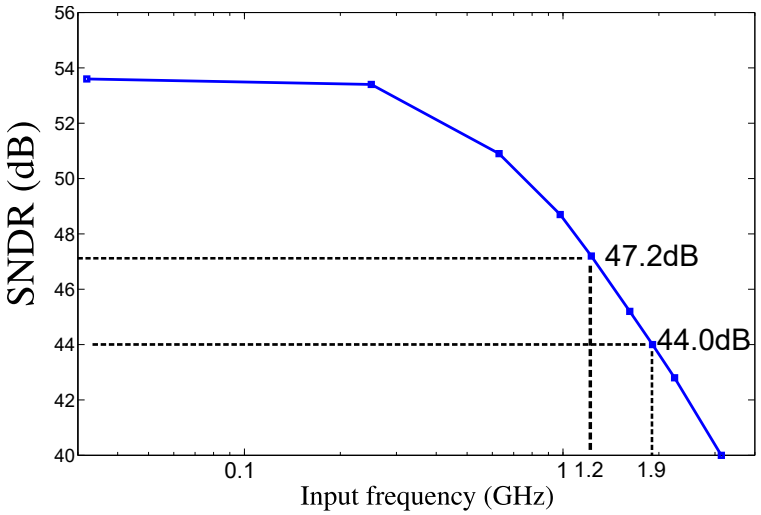


Figure 4.37: Effective bandwidth measurement, F_s=630MHz.

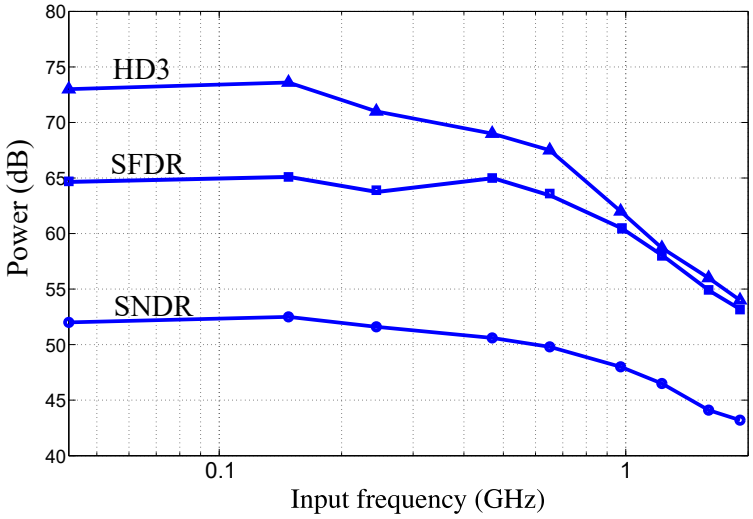


Figure 4.38: SNDR vs input frequency at F_s=2.5GHz.

ultimately set by requirements on jitter.

In the following sections, the causes of the unexpected performance degradations will be analyzed in details and corresponding solutions will be proposed.

4.9.1 V_{DD} reduction due to over-voltage limit

In measurement, when V_{DD} was 1.1V, transistor breakdown had been observed for the bottom plate sampling switches. The typical voltage rating of LVT transistors is 0.9V and the

	CICC 2015	ISSCC 2015	ISSCC 2015	ISSCC 2016	VLSI 2017	TCAS 2018	**This work**
Architecture	TI SAR	TI SAR	TI SAR	TI SAR	TI SAR	TI SAR	**TI SAR**
Technology(nm)	28	28	45	65	16	65	**28**
Resolution(bits)	10	10	10	10	10	10	**10**
F_S(GS/s)	5	5	1.7	2.6	2	2.3	**2.5**
F_{in}(GHz)	2	2.35	0.8	1.3	1	1.1	**1.2**
Supply	1	1.8/1.0	1.2	1.2	0.85	1.2/1.1	**1/0.9**
SNDR@DC	-	50	55	54.2	54.5	50.5	**52.0**
SNDR@Nyquist	41	46.2	51	50.6	50.6	47.6	**46.4**
Area(mm^2)	0.57	0.45	0.057	0.83	0.014	0.19	**0.3**
Power(mW)	76	150	15	18	10	31	**65**
FOM_{Walden}fJ/conv.	165	96	30	26	20	69	**151**
Skew calibration	Yes	Yes	Yes	Yes	Yes	No	**No**
Gain calibration	Yes	Yes	No	Yes	Yes	No	**No**
Offset calibration	Yes	Yes	Yes	Yes	Yes	Yes	**Yes**

Table 4.1: Performance comparison with state-of-art RF sampling ADCs.

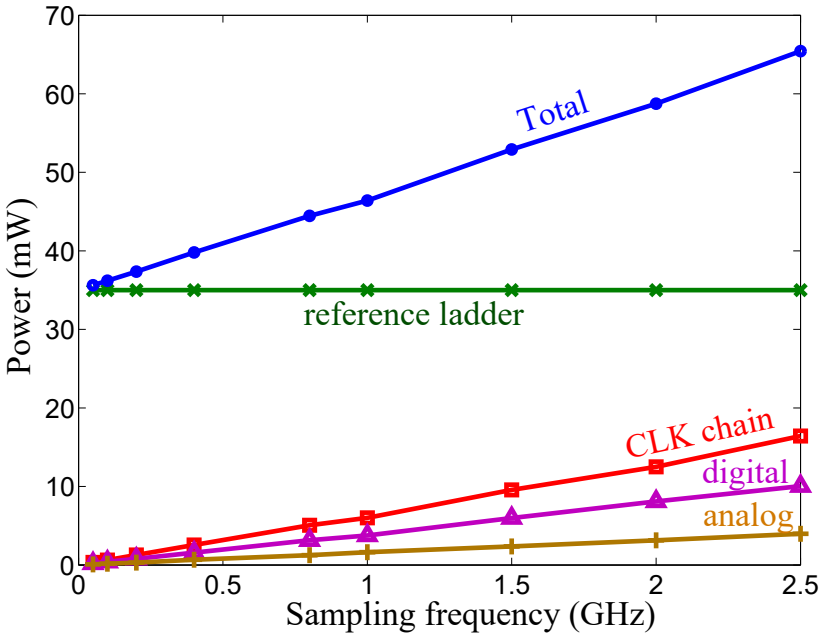


Figure 4.39: Power consumption versus sampling frequency.

maximum V_{GS} for safe operation is 1.1V. However, with bootstrapping circuits that raise gate voltages of bottom plate sampling switches to $2V_{DD}$, V_{GS} of these transistor would exceed 1.3V. For transistors to safely operate, the supply had been lowered to 0.9V in the later measurement. Back to the time of design, the voltage limit had been considered by me as a long term reliability concern rather than a hard limit that results in transistor breakdown, but it was indeed a strict limit. Bootstrapping should be applied with caution, especially in modern technology nodes with thin oxide and short channels. Breakdown would easily occur if safe operation requirements were not satisfied. Because of this supply reduction, some of the chip performance would unavoidably compromise, including the maximum sampling frequency, tracking bandwidth and harmonic distortion. These will be discussed in details in the following subsections.

4.9.2 Residual offsets

As shown in the low frequency spectrum measurement, interleaving spurs caused by offsets are larger than expected. Offsets also raise the noise floor as each comparison cycle is a 2bit flash comparison in 2bit/cycle SAR conversion. First we quantify the residual offsets

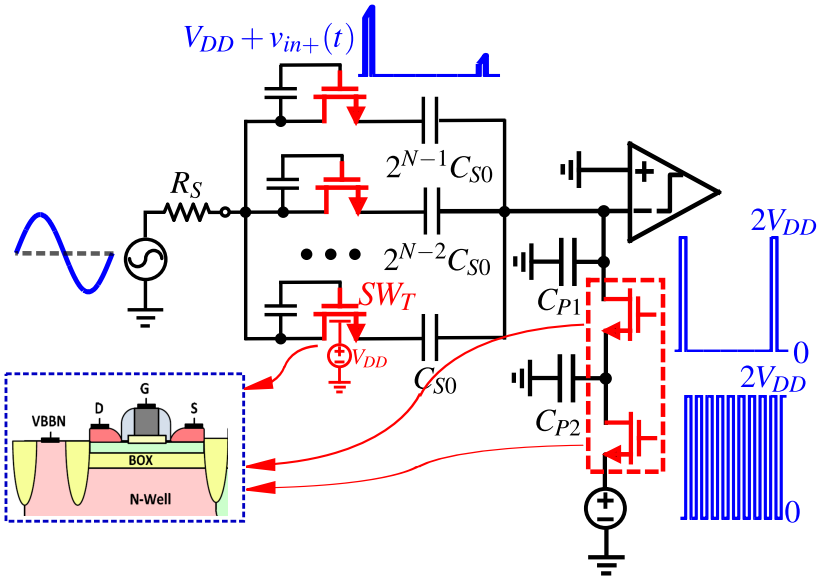
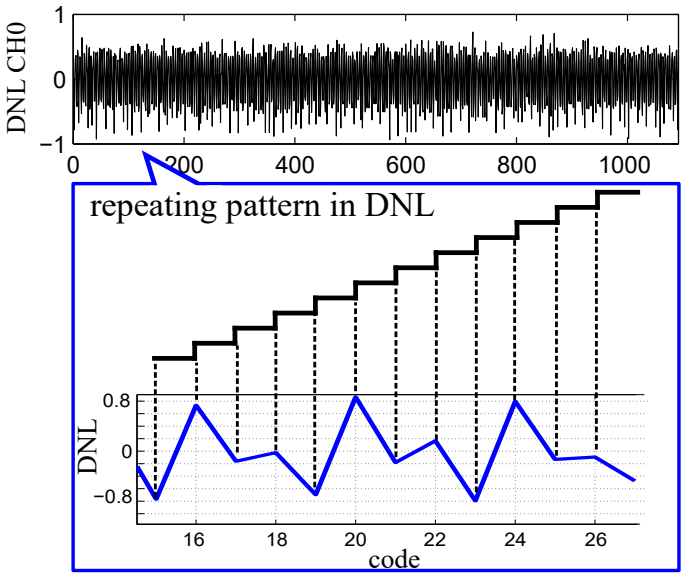


Figure 4.40: Bootstrapped bottom plate switches under large voltage stress.

from measurement by inspecting the measured DNL in each sub-channel. Fig.4.41 shows that the measured DNL of each sub-channel has a repeating pattern, from which the input referred offsets of all three comparators are extracted. Without noise, the measured DNL would display the exact repeating pattern. The extracted offsets have a standard variation of $\sigma_{V_{os}} = 0.52$mV with minimum/maximum offsets of -1.2mV/1.5mV from 24 comparators, 3 comparators per sub-channel.

The offset calibration is designed to suppress input referred offsets of the comparators to within $\pm$0.4mV after calibration. However, the measured residual offsets are larger by 3$\sim$4 times. This is caused by the different biasing conditions between comparators during actual conversion mode and comparators during calibration mode. The input voltages to the comparators are held on capacitors in CDACs during ADC conversion. They experience a common mode voltage step when CLK is strobed from low to high because of capacitive coupling paths labeled in Fig.4.42. M1,M2 extract charge from the capacitor array to establish the conducting channel when CLK becomes high. This common mode step is converted to a differential voltage that appears as an input offset [22] due to transistor mismatches, mainly V_t mismatches. However, during calibration mode, the comparator input terminals are connected to a DC voltage source, thus during calibration the voltage source supplies



$$\begin{cases} \text{DNL}_n = V_{os1} - 0 \\ \text{DNL}_{n+1} = V_{os2} - V_{os1} \\ \text{DNL}_{n+2} = V_{os3} - V_{os2} \\ \text{DNL}_{n+3} = 0 - V_{os3} \end{cases} \Rightarrow \text{solve}(V_{os1}, V_{os2}, V_{os3})$$

$\Downarrow$

	V_{os1}	V_{os2}	V_{os3}
CH0	1.5mV	0.45mV	-0.4mV
CH1	0.9mV	0.15mV	0.4mV
CH2	-0.3mV	-0.25mV	-0.7mV
CH3	-0.58mV	-0.8mV	0.68mV
CH4	-0.1mV	-0.4mV	-0.3mV
CH5	0.14mV	0.62mV	0.1mV
CH6	0.1mV	0.37mV	-1.2mV
CH7	0.7mV	0.17mV	-0.6mV

Figure 4.41: Calculate input referred offsets from measured DNL patterns.

the charge with no voltage step or conversion from common mode to differential mode. To further verify the cause, we calculate the residual offsets using mismatch coefficients from the PDK. V_t mismatch between the comparator input pair is

$$\sigma_{\Delta V_t} = \frac{A_{V_t}}{\sqrt{W \times L}} = \frac{2\text{mV} \cdot \mu\text{m}}{\sqrt{2\mu\text{m} \times 0.045\mu\text{m}}} = 6.7\text{mV} \tag{4.27}$$

The common-mode step (common-mode kick back) is obtained from simulation,

$$\Delta V_{in,CM} = \frac{C_{OX}}{C_S}(V_{GS} - V_{t1,2}) = 20\text{mV}, V_{GS} - V_t = 550 - 300 = 250(\text{mV}) \tag{4.28}$$

The residual offsets can be then estimated

$$\sigma_{V_{OS,res}} = \frac{C_{OX}}{C_S}\sigma_{\Delta V_{t1,2}} = \mathbf{0.54mV} \tag{4.29}$$

This matches the measured residual offsets whose σ is 0.52mV. From Matlab simulation where input referred offsets are manually inserted, the residual offsets limit the total SNDR to 54dB in an otherwise ideal interleaved ADC. This is the most significant performance limit for measurements at low input frequencies. A proper design should calibrate when the comparators are operating in the same condition as in the actual ADC conversion, that is, driven at the inputs by capacitor arrays. This will absorb the residual offsets into calibration and suppress offsets to the designed range.

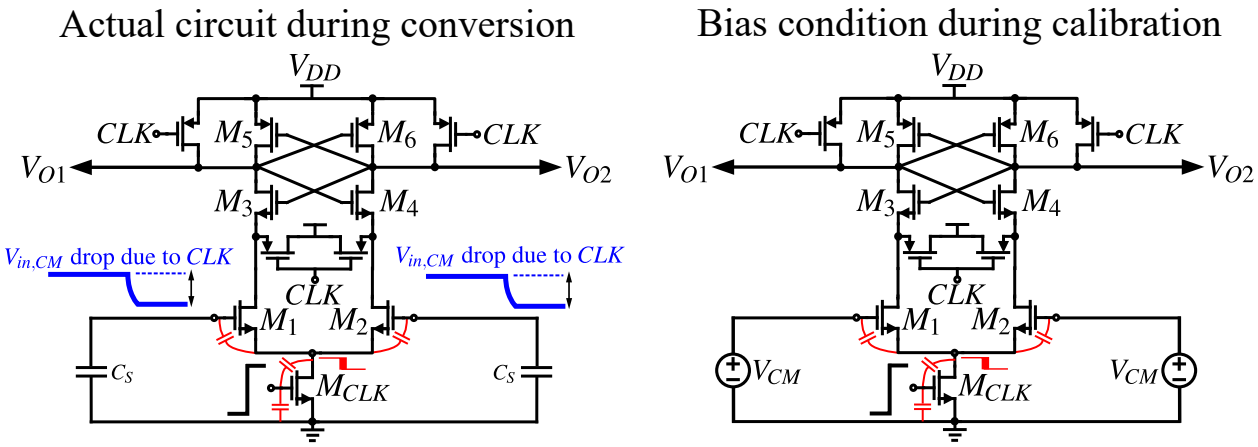


Figure 4.42: Comparator during actual ADC conversion and calibration mode.

4.9.3 Gain mismatch due to reference mismatch

The measured low frequency spectrum also displays interleaving spurs at $F_s/8 \pm f_{in}$, $2F_s/8 \pm f_{in}$ ... which can be caused only by gain mismatches. The impact of timing skew/bandwidth mismatch is negligible at low frequencies. Gain mismatches arise from reference mismatches as shown in Fig.4.43. The decoupling capacitors are all placed to the reference ladder while sub-channels do not have dedicated local decoupling capacitors. Because all sub-channels pull charge asynchronously at high speeds, 15GHz(2.5GHz×6) at the maximum sampling frequency, it is difficult to estimate actual voltage fluctuations on the reference line. The resistive wiring results in a systematic voltage drop along the wire for reference distribution and induces interleaving spurs. The currents through the wires are not DC, but pulses. From measured spectrum with low frequency input, the average spur level due to reference mismatch is around -65dBc, we estimate the reference mismatch from the measurement

$$\sigma_{\Delta V_{ref}} = V_{ref} \times 10^{-\frac{65}{20}} = 0.5\text{mV} \tag{4.30}$$

0.5mV is well below 1LSb, but its impact is already noticeable. It limits the total SNDR to 59dB, obtained from Matlab simulation where static(DC) reference mismatch are inserted in an otherwise overall ADC. The improved design would place decoupling capacitors at the reference ladder in each sub-channel ADC so that the overall fluctuation on the reference line become quiet. Reference mismatch is most severe during MSB cycles, so decoupling

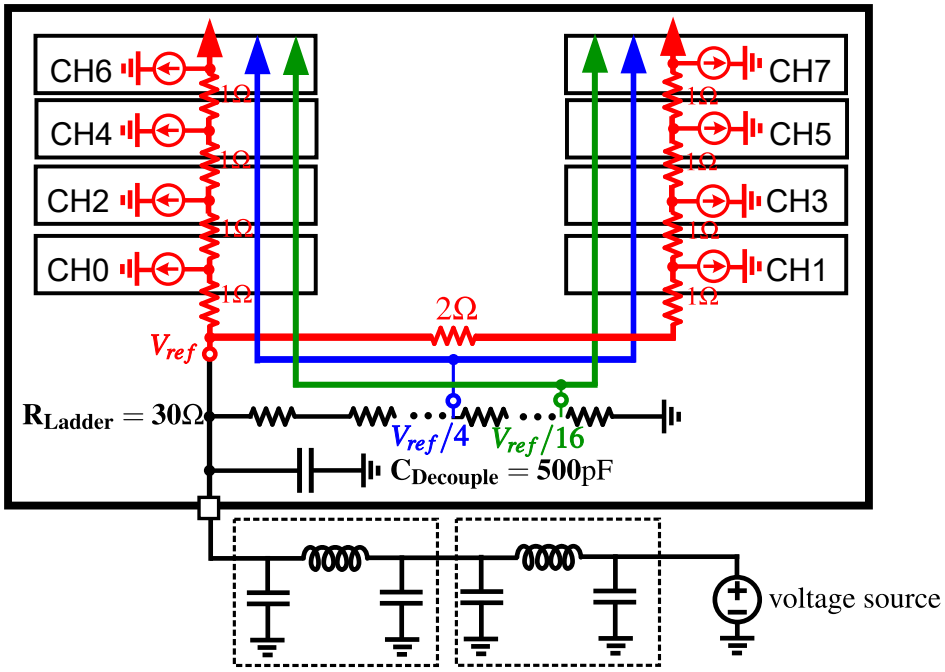


Figure 4.43: Reference mismatch caused by resistive routing and cub-channels pulling currents asynchronously.

capacitors should be placed mainly to the full scale reference V_{ref}. Fractional reference voltages serve LSBs, where sensitivity to reference mismatch is reduced proportionally. Cross coupling among different channels are also minimized in this arrangement. Furthermore, the power consumption of the reference ladder is significantly higher than necessary. Since fractional references $\frac{V_{ref}}{4}$, $\frac{V_{ref}}{16}$ are used only during LSB conversions, the currents these two nodes need to support are much lower than MSB conversions. Currently the reference ladder alone consumes 35mW static power. Reducing the reference power by 80% to 5.6mW will not limit the conversion speed as shown in Fig.4.44.

4.9.4 Timing skew/bandwidth mismatch revisited

The measured spectrum with high frequency input displays spurs at $F_s/8 \pm f_{in}$, $2F_s/8 \pm f_{in}$ ... that increase with input frequencies. These are caused by timing skew and/or bandwidth mismatches. However, these two cannot be distinguished from each other from measurement because bandwidth mismatch is equivalent to timing skew to the first order. Since spurs

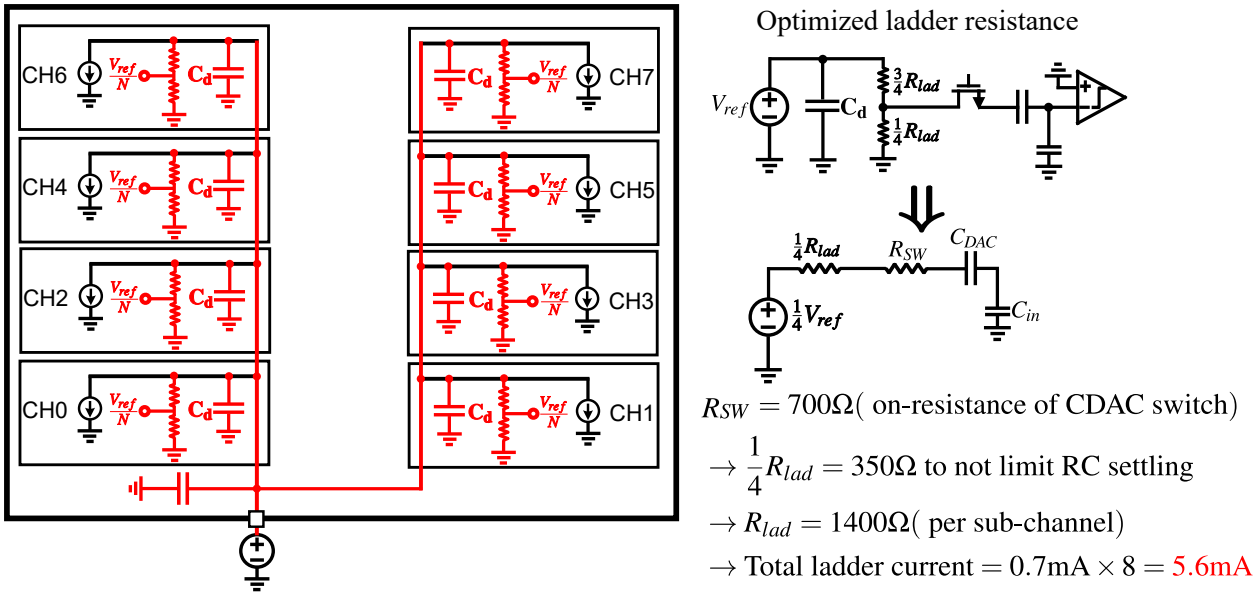


Figure 4.44: Improved reference ladder design and reference distribution scheme.

caused by timing skew/bandwidth mismatches appear at the same frequencies as spurs caused by reference(gain) mismatches, we isolate them by observing FFT phases of each sub-channel. Without timing skew/bandwidth mismatches, adjacent sub-channels have a uniform phase difference of

$$\Delta\Phi = 2 \times \pi \times \underbrace{8}_{\text{No. of TI}} \times \underbrace{\text{Deci}}_{\text{Decim. ratio}} \times f_{in} \times \underbrace{T_s}_{1/F_s} \tag{4.31}$$

Now with timing skew/bandwidth mismatches, phase mismatches appear among sub-channels

$$\Delta\phi = 2 \times \pi \times 8 \times \text{Deci} \times f_{in} \times \underbrace{\Delta t}_{\text{skew}} + \underbrace{\frac{\Delta f_0}{f_0} \cdot \frac{f_{in}}{f_0}}_{\text{BW mismatch}} \tag{4.32}$$

From measurement, phase mismatches with Nyquist 1.2GHz input are calculated

$$\begin{aligned} \Delta\phi &= [1.9,\ \ 0.2,\ \ 1.6,\ -0.7,\ -0.8,\ \ 2.1,\ -1.7] \times 10^{-3}\text{rad} \\ \sigma_{\Delta\phi} &= \mathbf{1.4 \times 10^{-3} rad} \end{aligned} \tag{4.33}$$

This amount of phase mismatch would limit the SNDR of the interleaved ADC to 52dB. This is worse than the impact of offsets and is the most significant degradation of performance at Nyquist frequency.

The spurs caused by bandwidth mismatches grow with the input frequency and is inverse proportional to the tracking bandwidth. The tracking bandwidth in the actual chip degrades for two reasons. First, because of V_{DD} reduction from 1.1V to 0.9V, the on-resistances of the sampling switches increase. Second, layout parasitics, mainly capacitors, were underestimated and there was no time to extract them before the tape-out deadline. The parasitic capacitors from long routing wires connecting input signals to each sub-channel pose additional penalty on tracking bandwidth as shown in Fig.4.45. These routing wires are on top metal layer with a total resistance less than 5Ω. The actual tracking bandwidth is estimated to be 2.7GHz with simple calculation while simulation reveals it to be 2.4GHz.

With the same β mismatch, bandwidth mismatch increases to 0.2% as V_{DD} reduction raises the contribution of the smallest sampling switch to the total resistance

$$\sigma_{\Delta f_{0,\beta}} = 0.8\% \times \frac{33\Omega}{129\Omega} = 0.2\% \tag{4.34}$$

Additional, V_t mismatch needs to be included as V_{DD} reduction results in reduced V_{OV}, which amplifies the impact of V_t mismatch.

$$\begin{aligned} \sigma_{\Delta V_t} &= \frac{A_{V_t}}{\sqrt{W \times L}} = \frac{2\text{mV} \cdot \mu\text{m}}{\sqrt{8\mu\text{m} \times 0.03\mu\text{m}}} = 4\text{mV} \\ \sigma_{\Delta f_{0,V_t}} &= \frac{\sigma_{\Delta V_t}}{V_{GS} - V_t} \times \frac{33\Omega}{129\Omega} = \frac{2.7\text{mV}}{0.6\text{V}} \times \frac{33\Omega}{129\Omega} = 0.1\% \end{aligned} \tag{4.35}$$

Thus, the net bandwidth mismatch is $\sigma_{\Delta f_0}/f_0 = 0.23\%$. It is converted to phase mismatch of

$$\sigma_{\Delta\phi_{BW}} = 0.23\% \times \frac{1.22\text{GHz}}{2.4\text{GHz}} = \mathbf{1.2 \times 10^{-3} rad} \tag{4.36}$$

The reduced V_{DD} also slows down the sampling clock and results in worsened timing skew

$$\begin{aligned} \sigma_{\Delta t} &= \frac{\sigma_{\Delta V_t}}{\mathrm{d}V/\mathrm{d}t} = \frac{2.7\text{mV}}{1.2\text{V}/35\text{ps}} = \mathbf{78fs} \\ \Rightarrow \sigma_{\Delta\phi_{skew}} &= 2\pi \times f_{in} \times \sigma_{\Delta t} = 2\pi \times 1.22\text{GHz} \times 78\text{fs} = \mathbf{0.6 \times 10^{-3} rad} \end{aligned}$$

The overall phase mismatch due to timing skew/bandwidth mismatches is then

$$\sigma_{\Delta\phi} = \sqrt{\langle \Delta\phi_{skew}^2 \rangle + \langle \Delta\phi_{BW}^2 \rangle} = \mathbf{1.3 \times 10^{-3} rad} \tag{4.37}$$

The calculated phase mismatch is close to measurement and shows clearly the high frequency performance degradation.

$$f_0 = \frac{1}{2\pi[(R_S + R_{wire})C_{wire} + \Sigma R \times (C_{wire} + C_S)]}$$
$$= \frac{1}{2\pi} \times (55\Omega \times 100\text{fF} + 125\Omega \times 420\text{fF})^{-1} = 2.7\text{GHz}$$

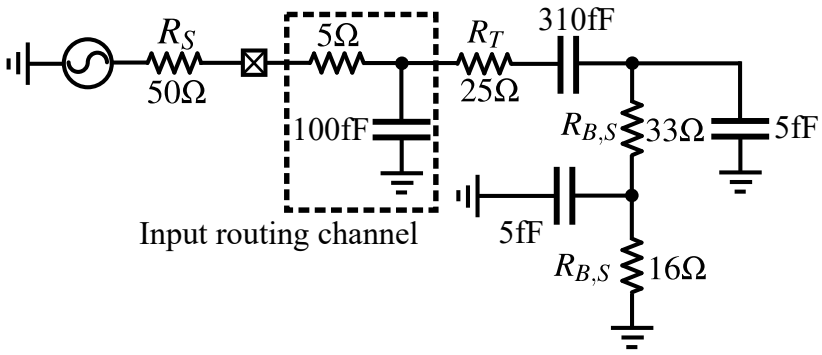


Figure 4.45: Simplified equivalent circuits with increased transistor R_{on} and additional parasitics from layout.

There are two solutions to suppress the impacts of timing skew/bandwidth mismatches. The first option is to better control the layout and re-optimize the transistor size to balance HD3 and bandwidth mismatches. The current transistor sizing is optimized to primarily minimize harmonic distortions without enough consideration for suppressing bandwidth mismatch. An improved design would balance the two, which will be described in details in the following section. The second option is to introduce timing calibration circuits, usually implemented with a programmable delay line that fine tunes the sampling instants of each sub-channel ADC. In modern CMOS technology node, the resolution of the delay line can be designed to be less than 50fs. This is well enough for 10bit resolution ADCs with sampling frequencies up to 5GHz. This resolution of the delay line is mainly limited by power consumption. It is possible to achieve finer resolutions, but at the cost of additional power. Applying timing calibration circuits is a more sophisticated and complete solution, but it requires additional hardware in DSP for error detection and the delay lines with fine resolution also consume significant power.

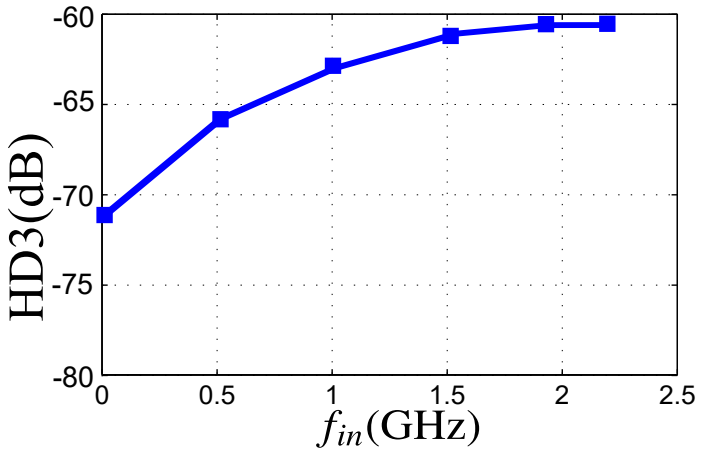


Figure 4.46: HD3 of re-optimized passive master/slave T/H, top plate switch size: 48μm/28nm, bottom plate slave switch size: 32μm/28nm, bottom plate master switch size: 48μm/28nm.

4.9.5 Harmonic distortion revisited

The harmonic distortions degrade for the previously mentioned reason of V_{DD} reduction. The measured harmonic distortion is dominated by HD3 and is around 10dB worse than designed. V_{DD} reduction leads to smaller transistor over-drive voltages of the sampling transistors [40], which is the main cause of observed degradations in harmonic distortions.

$$\begin{aligned} V_{OV,design} &= 1.8\text{V} - 0.5\text{V} - 0.3\text{V} = 1.1\text{V} \\ V_{OV,actual} &= 1.4\text{V} - 0.5\text{V} - 0.3\text{V} = 0.7\text{V} \\ \text{HD3}_{R_{on}} &\sim \frac{1}{V_{OV}^3} \rightarrow \text{Degradation} = 20 \times \log_{10}((1.1/0.7)^3) = 12\text{dB} \end{aligned} \tag{4.38}$$

At Nyquist input frequency 1.2GHz, the total harmonic distortion limits the SNDR to 58dB, which is still below spurs due to bandwidth mismatches. An improved design would redesign the T/H under 0.9V supplies and balance the impact of harmonic distortion and bandwidth mismatches. Fig.4.46 plots the simulated HD3 in a resized T/H at 0.9V supply voltage. Bottom plate switches are enlarged to reduce mismatch at the cost of higher HD3. Now the bandwidth mismatch is reduced by 2 times with 4 times larger switches, which reduces spurs due to bandwidth mismatches by 2 times, improving the SNDR set by bandwidth mismatches from -52dB to -58dB in an otherwise same circuits. Further improvements are

achievable with better layouts that reduce routing capacitors. HD3 is still below -58dB in the resized circuit and does not limit the overall SNDR.

4.9.6 Sampling jitter

The measured noise floor grows with input frequencies Fig.4.47 and this is caused by the sampling jitter. At high frequencies the roll-off is around 4dB/octave instead of 6dB/octave, which is caused by the additional contribution from offsets to the noise floor. From measurement, the sampling jitter is extracted to be 130fs by subtracting contribution of offsets, which is still higher than estimated during design. The degradation is mainly caused by layout parasitic capacitors that had not been captured during schematic design.

As shown in Fig.4.48, a total of eight stages of buffers are in the clock chain that converts the sinewave input clock to a square wave clock and distribute it to all sub-channels. The output jitter arising from thermal noise within the clock buffer chain can be readily calculated using the established noise calculation in [22, 90]. The noise currents from each buffer stage integrate onto each stage's loading capacitor when clock transitions, and the noise voltages accumulated on the capacitors are converted to jitter in time domain by dividing the noise voltage by the slope of the clocks.

$$\begin{aligned} \text{Noise voltage: } \langle v_{out}^2 \rangle &= 4\text{kT}\gamma G_m \times \frac{t_r}{C_L^2} \text{ ([22, 90])} \\ \text{Clock slope: } \frac{\mathrm{d}v_{out}}{\mathrm{d}t} &= \frac{G_m \cdot \frac{V_{DD}}{2}}{C_L} \\ \Rightarrow \text{Output jitter: } \langle \Delta t^2 \rangle &= \frac{\langle v_{out}^2 \rangle}{\left(\mathrm{d}v_{out}/\mathrm{d}t\right)^2} = 16\text{kT}\gamma \cdot \frac{1}{G_m}\frac{t_r}{V_{DD}^2} \end{aligned} \tag{4.39}$$

Using simulated transconductances and clock transition times of each stage in transient simulation, the total output jitter of the sampling clock is obtained.

$$\langle \Delta t_{sample}^2 \rangle = \sum_{i=1}^{8} \langle \Delta t_i^2 \rangle \;\rightarrow\; \sigma_{\Delta t_{sample}} = \mathbf{77fs} (< \text{measurement}) \tag{4.40}$$

PSS/Pnoise simulation at schematic level shows a total output jitter of 80fs, which is close to calculated. The most significant noise contribution is from the first stage, where the input is a sinewave that has the longest window for noise to accumulate, and the second

largest contribution is from the last stage, which has the heaviest loading capacitors. Now we rewrite the jitter expression to investigate the impact of layout parasitics,

$$\begin{aligned} \langle \Delta t^2 \rangle &= 16\mathrm{kT}\gamma \cdot \frac{1}{G_m} \frac{t_r}{V_{DD}^2} \\ \text{with } t_r &= \frac{C_L}{\frac{1}{2}G_m} \text{ ([90])} \\ \Rightarrow \langle \Delta t^2 \rangle &= 16\mathrm{kT}\gamma \cdot \frac{1}{G_m} \frac{1}{V_{DD}^2} \frac{C_L}{\frac{1}{2}G_m} \\ &= 32\mathrm{kT}\gamma \cdot \frac{C_L}{(G_m V_{DD})^2} \end{aligned} \tag{4.41}$$

This shows clearly how additional parasitic capacitors directly degrade jitter. Using the same calculation and updated rise/fall times obtained from post-layout simulations, the jitter increases to 105fs. Parasitic resistors do not directly contribute jitter as they appear in series with the loading capacitor, but they slow down the sampling clock, allowing longer time windows for noise to accumulate and resulting in degraded jitter. With additional 80fs jitter from the instrument, the total sampling jitter is 135fs, which is close to measurement. The sampling jitter limits the SNDR to 60dB.

$$\mathrm{SNR}_{\mathrm{jitter}} = 20\log_{10}\left(\frac{1}{2\pi \times 1.2\mathrm{GHz} \times 135\mathrm{fs}}\right) = 60\mathrm{dB} \tag{4.42}$$

The impact of sampling jitter is noticeable, but remains below contributions of bandwidth mismatches and harmonic distortions to the total SNDR.

4.9.7 Imperfections breakdown

Fig.4.49 plots the breakdown of various imperfections versus the input frequency at the sampling frequency of 2.5GHz. At low frequencies, comparator residual offsets are the most significant performance degradation source followed by reference mismatches. Input referred noise across all frequencies falls below the quantization noise floor. At high frequencies, bandwidth mismatch induced spurs dominates over other imperfections. Harmonic distortion is the second largest frequency dependent degradation at high frequencies. This highlights the importance of controlling bandwidth mismatch and maintaining a healthy tracking bandwidth. In this design, with three switches in series, achieving high bandwidth becomes more

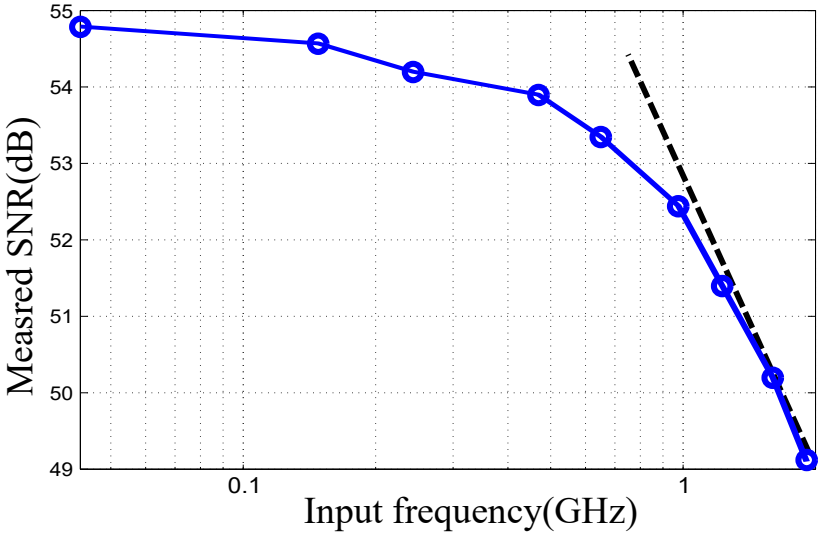


Figure 4.47: SNR versus input frequency to extract sampling jitter.

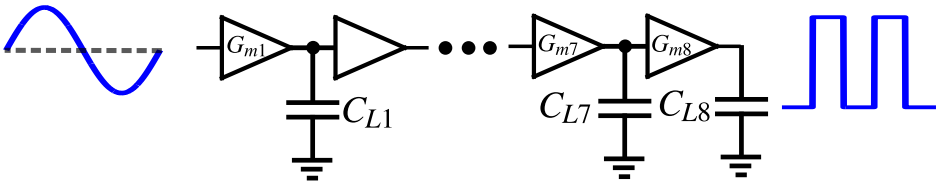


Figure 4.48: Clock buffer chain.

difficult than other designs where T/H circuits involve only one or two transistors. Sampling jitter and timing skew are still well controlled in the fabricated chip and do not pose significant limit to the overall performance.

4.9.8 Future improvements

The fabricated chip has demonstrated the potential of high-frequency moderate-resolution ADC design driven by complete circuit understandings. With minimum digital calibration, only comparator offsets calibrated in foreground, the ADC is still able to achieve comparable SNDR with state-of-arts. Due to lack of time, some of the design choices could not be fully optimized. Future improvements, as described in the previous section, would include

- Absorb common-mode kickback into offset calibration mode;
- Replace standard cells used in SAR logic with custom logic circuits;

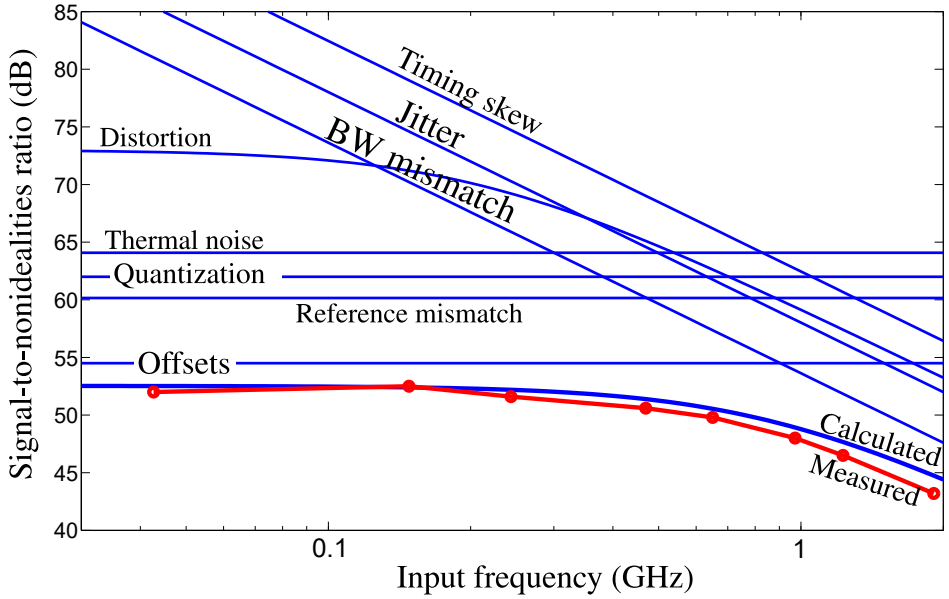


Figure 4.49: Calculated circuit imperfections breakdown.

- Reduce power consumption in the reference ladder and distribute the ladders to each sub-channel;
- Place local decoupling capacitors in each sub-channel;
- Re-size T/H to balance the impacts between harmonic distortion and bandwidth mismatches;
- Improve the layout. The passive T/H and SAR conversion speed will both benefit from a improved layout design.

Projected performance improvements and power reduction are shown in Fig.4.50 and Fig.4.51. Offsets would be reduced to the design boundaries of $\pm 0.4mV$. 4dB SNDR improvement can be achieved from correct offset calibration. Bandwidth mismatch is reduced by 2 times in the re-sized T/H as described in section 4.9.5. The improvement at high input frequencies also assumes a better layout that would maintain the tracking bandwidth around 5GHz, which is feasible given more time for implementation.

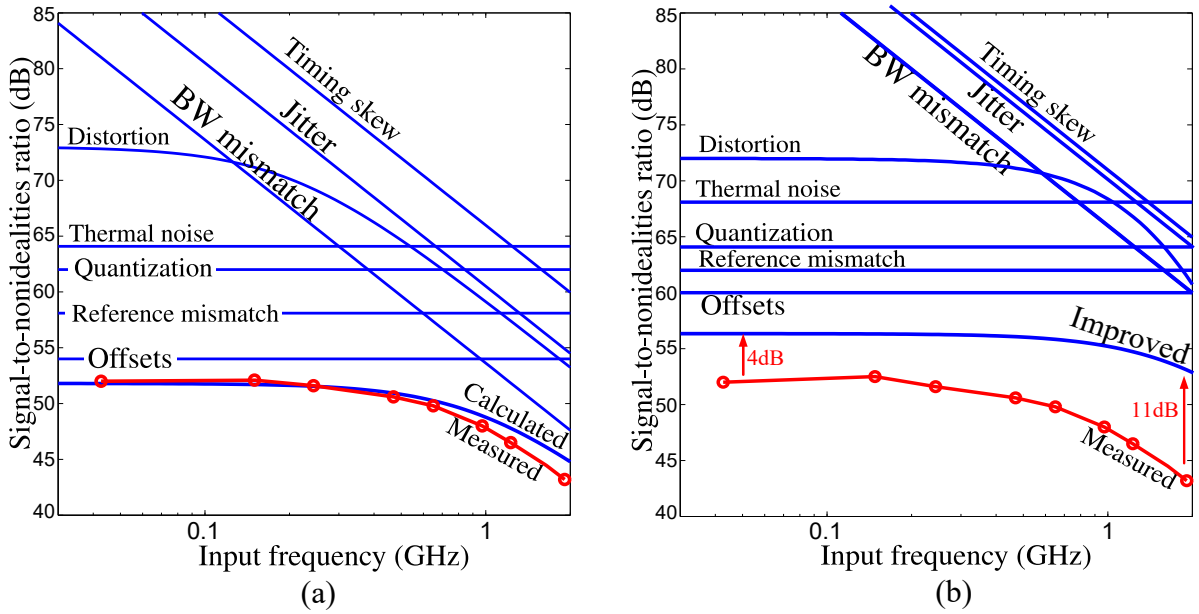


Figure 4.50: Projected performance improvement compared with measurement.

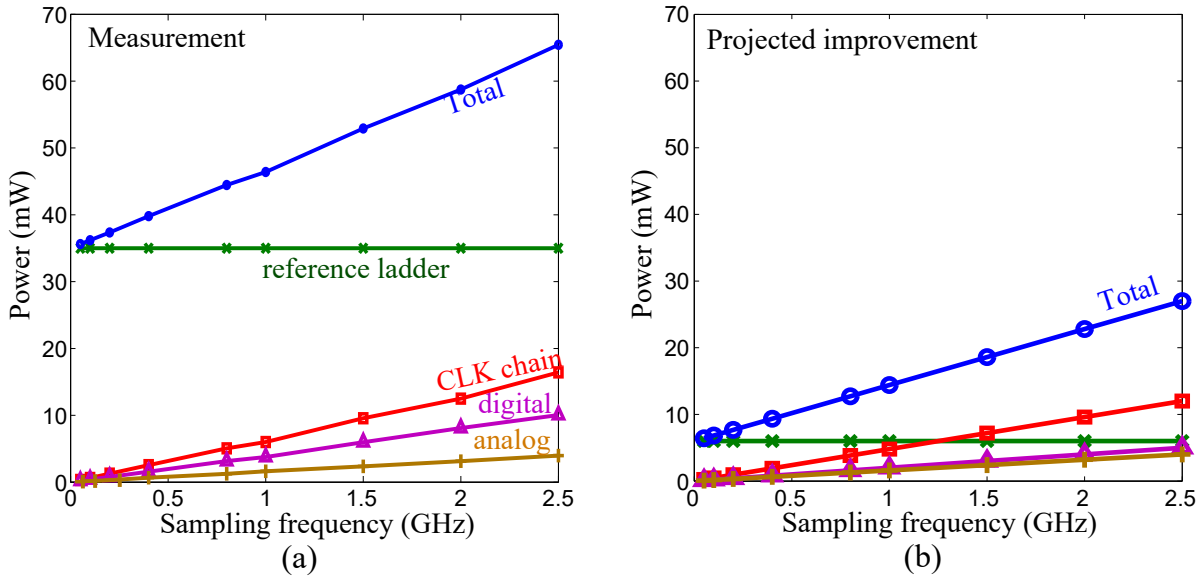


Figure 4.51: Projected power reduction compared with measurement.

4.10 Conclusion

This dissertation has demonstrated the potential of designing high speed ADCs based on complete circuit understandings. Limitations have been quantitatively studied and compared with measurement. While digital calibration is a powerful tool to combat many imperfections

in analog circuits, it is not always easy to implement and poses additional hardware overhead by itself. It should be only used to further improve performance of a fully optimized design rather than a universal solution that rescues any poorly designed circuits. In the regime of below 5GHz sampling speed and 10bit resolution, as this dissertation reveals, a fully optimized simple design is able to achieve the desired performance without intensive digital calibration. The analytical study presented in this dissertation provides a frame that enables designers to have a good first-cut estimate for high speed ADC designs and make correct design choices early on. It is good to have digital calibration, but only when necessary and done properly.

CHAPTER 5

Conclusions

This dissertation has shown that how modern mixed-signal circuit design can benefit from circuit fundamental understandings. Designing comparators, ADCs or PLLs are certainly different from designing opamps, which has a well established framework that contains all aspects of design optimization. There is no ambiguity in opamp design today. This dissertation contributes to building similar frameworks for mixed-signal circuit design that explain existing design techniques and enables more optimized designs. This subject as a whole is of course beyond one single thesis, but this dissertation has proved that mixed-signal circuit design, which often involves time-varying systems with hard non-linearity, is not a field where the unknown prevails understanding. These circuits/systems can be analytically characterized and understood with proper approaches. And finally, they can be designed with knowledge rather than trial and error.

Key contributions of this dissertation are summarized as below.

For the comparators,

- For the first time, dynamic offsets in regenerative latch due to common mode to differential mode conversion is clearly visualized and explained using phase plane plots.
- For the first time, the operations of the strongARM comparator and various double-tail comparators are clearly defined. Their operation has been condensed to the internal dynamic amplifiers that ultimately determine offsets and noise, which are then quantitatively studied and verified.
- For the first time, closed-form expressions for offsets arising from different mismatches in strongARM comparators are derived. This answers the question if these offsets are

VDD dependent and whether they can be reliably calibrated.

- For the first time, the design trade-off between a strongARM comparator and a double-tail comparator is clearly defined. This extends to all regenerative comparators.
- Design guidelines for regenerative comparators are provided.

For bang-bang PLLs,

- For the first time, bang-bang PLLs are characterized based on linearized frequency domain analysis that renders closed-form expressions for loop gain, bandwidth, phase noise and jitter.
- Optimum loop gain to minimize jitter of bang-bang PLLs is derived using practical design parameters. Calibration to find such optimum is then explained using the established frequency domain analysis.
- Simple design guidelines for bang-bang PLLs are provided.

For ADC design,

- The prototype ADC is an example to demonstrate the power of the systematic design approach of ADCs. Trade-offs are quantitatively compared. Imperfections are characterized and budgeted at the phase of design.
- As simple as it is, the prototype ADC performs fairly well (with degradations not fully captured by design) compared with stat-of-arts that have intensive calibration, but in a considerably shorter time frame and with much less complexity. This shows the power of optimization driven by circuit understandings.
- Passive T/H is explored and implemented in the prototype ADC. Its potential has been clearly demonstrated. With the analytical approach presented in this dissertation, designers can determine early on if a passive T/H is sufficient for certain application or more sophisticated circuits would be necessary.

Because of lack of time, the prototype ADC has not fully reached its potential and it can be further improved in many perspectives. Some are immediate steps to fix the observed degradations as below.

- Absorb common-mode kickback into offset calibration mode;
- Replace standard cells used in SAR logic with custom logic circuits;
- Reduce power consumption in the reference ladder and distribute the ladders into each sub-channel;
- Place decoupling capacitors locally in each sub-channel;
- Re-size T/H to balance the impacts between harmonic distortion and bandwidth mismatches;
- Improve layout. The passive T/H and SAR conversion speed will both benefit from improved layout design.

Yet, some questions deserve more study and as mentioned above, they need to be analytically addressed and included in a larger framework.

- 2bit/cycle SAR is implemented in this design to reduce number of interleaving channels. Its poses stringent requirements on comparators. But it is also feasible to speed up 1bit/cycle SAR with other CDAC switching scheme and potentially achieve comparable speed improvement as a 2bit/cycle SAR.
- Hybrid architectures can become valuable. For example, a pipelined SAR ADC improves conversion speed with the aid of an active amplifier. Though it results in additional power and design efforts, it is possible to achieve a better design trade-off than a pure SAR ADC.
- With master/slave T/H, timing skews are suppressed with a global master clock. The cost is the additional transistor in series, which actually limits bandwidth matching

and distortion. Is it possible to implement other clocking scheme that controls timing skew without increasing number of switches in T/H?

- In this design, no calibration is introduced except offset calibration. It is always an open question about the overhead from calibration. If it is implemented efficiently, it may significantly improve the performance with possibly little overhead. But it needs much more consideration in both analog and digital circuit, and the partitioning between analog and digital domains is always challenging. This is where significant contribution can be made to the whole ADC design community.

REFERENCES

[1] B. Wicht, T. Nirschl, and D. Schmitt-Landsiedel, "Yield and speed optimization of a latch-type voltage sense amplifier," *IEEE J. of Solid-State Circuits*, vol. 39, no. 7, pp. 1148–1158, 2004.

[2] L. Kull, T. Toifl, M. Schmatz, P. Francese, C. Menolfi, M. Brandli, M. Kossel, T. Morf, T. Andersen, and Y. Leblebici, "A 3.1 mw 8b 1.2 gs/s single-channel asynchronous sar adc with alternate comparators for enhanced speed in 32 nm digital soi cmos," *Solid-State Circuits, IEEE Journal of*, vol. PP, no. 99, pp. 1–10, 2013.

[3] D. Schinkel, E. Mensink, E. Klumperink, E. van Tuijl, and B. Nauta, "A Double-Tail Latch-Type Voltage Sense Amplifier with 18ps Setup+Hold Time," in *IEEE International Solid-State Circuits Conference. Digest of Technical Papers*, Feb 2007, pp. 314–605.

[4] M. Miyahara and A. Matsuzawa, "A low-offset latched comparator using zero-static power dynamic offset cancellation technique," in *2009 IEEE Asian Solid-State Circuits Conference*, Taipei, Taiwan, Nov 2009, pp. 233–236.

[5] S. Jang, S. Kim, S. Chu, G. Jeong, Y. Kim, and D. Jeong, "An Optimum Loop Gain Tracking All-Digital PLL Using Autocorrelation of Bang-Bang Phase-Frequency Detection," *IEEE Trans. Circuits Syst. II*, vol. 62, no. 9, pp. 836–840, 2015.

[6] A. Vareljian, M. Moussavi, W. Bereza, W. Fergusson, C. Berndt, and R. Patel, "Non-linear Behavior Study in Digital Bang-Bang PLL," in *IEEE CICC Proc.*, Sept 2009, pp. 339–342.

[7] S. Levantino, G. Marzin, C. Samori, and A. Lacaita, "A Wideband Fractional-N PLL With Suppressed Charge-Pump Noise and Automatic Loop Filter Calibration," *IEEE J. Solid-State Circuits*, vol. 48, no. 10, pp. 2419–2429, Oct 2013.

[8] D. Murphy, J. Rael, and A. A. Abidi, "Phase Noise in LC Oscillators: A Phasor-Based Analysis of a General Result and Loaded Q," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 57, no. 6, pp. 1187–1203, June 2010.

[9] A. Rylyakov, J. Tierno, H. Ainspan, J.-O. Plouchart, J. Bulzacchelli, Z. Deniz, and D. Friedman, "Bang-bang Digital PLLs at 11 and 20GHz with Sub-200fs Integrated Jitter for High-Speed Serial Communication Applications," in *IEEE ISSCC Dig.*, Feb 2009, pp. 94–95,95a.

[10] T. Kuan and S. Liu, "A Digital Bang-Bang Phase-Locked Loop With Automatic Loop Gain Control and Loop Latency Reduction," in *IEEE Symp. on VLSI Circuits (VLSIC)*, June 2015, pp. C138–C139.

[11] D. Tasca, M. Zanuso, G. Marzin, S. Levantino, C. Samori, and A. Lacaita, "A 2.9 to 4.0-GHz Fractional-N Digital PLL With Bang-Bang Phase Detector and 560-fs RMS

Integrated Jitter at 4.5-mW Power," *IEEE J. of Solid-State Circuits*, vol. 46, no. 12, pp. 2745–2758, Dec 2011.

[12] T. H. O'Dell, *Circuits for Electronic Instrumentation.* Cambridge, UK: Cambridge U. P., 1991.

[13] R. Sarpeshkar, J. Wyatt, J. L., N. C. Lu, and P. D. Gerber, "Mismatch sensitivity of a simultaneously latched CMOS sense amplifier," *Solid-State Circuits, IEEE Journal of*, vol. 26, no. 10, pp. 1413–1422, 1991.

[14] R. Kraus, "Analysis and reduction of sense-amplifier offset," *Solid-State Circuits, IEEE Journal of*, vol. 24, no. 4, pp. 1028–1033, 1989.

[15] A. Nikoozadeh and B. Murmann, "An analysis of latch comparator offset due to load capacitor mismatch," *IEEE Trans. on Circuits and Systems II*, vol. 53, no. 12, pp. 1398–1402, 2006.

[16] T. Kobayashi, K. Nogami, T. Shirotori, and Y. Fujimoto, "A current-controlled latch sense amplifier and a static power-saving input buffer for low-power architecture," *IEEE J. of Solid-State Circuits*, vol. 28, no. 4, pp. 523–527, 1993.

[17] L. Sumanen, M. Waltari, and K. Halonen, "A mismatch insensitive cmos dynamic comparator for pipeline a/d converters," in *Electronics, Circuits and Systems, 2000. ICECS 2000. The 7th IEEE International Conference on*, vol. 1, 2000, pp. 32–35 vol.1.

[18] J. He, S. Zhan, D. Chen, and R. L. Geiger, "Analyses of Static and Dynamic Random Offset Voltages in Dynamic Comparators," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 56, no. 5, pp. 911–919, 2009.

[19] J. Kim, B. S. Leibowitz, J. Ren, and C. J. Madden, "Simulation and analysis of random decision errors in clocked comparators," *IEEE Trans. on Circuits and Systems I*, vol. 56, no. 8, pp. 1844–1857, 2009.

[20] P. Nuzzo, F. De Bernardinis, P. Terreni, and G. Van der Plas, "Noise analysis of regenerative comparators for reconfigurable ADC architectures," *IEEE Trans. on Circuits and Systems I*, vol. 55, no. 6, pp. 1441–1454, 2008.

[21] P. Nuzzo, C. Nani, C. Armiento, A. Sangiovanni-Vincentelli, J. Craninckx, and G. Van der Plas, "A 6-Bit 50-MS/s threshold configuring SAR ADC in 90-nm digital CMOS," *IEEE Trans. on Circuits and Systems I*, vol. 59, no. 1, pp. 80–92, 2012.

[22] A. Abidi and H. Xu, "Understanding the Regenerative Comparator Circuit," in *Custom Integrated Circuits Conference*, San Jose, CA, Sept 2014, pp. 1–8.

[23] Y. Asada, K. Yoshihara, T. Urano, M. Miyahara, and A. Matsuzawa, "A 6bit,7mW,250fJ,700Ms/s Subranging ADC," in *Solid-State Circuits Conference, 2009. A-SSCC 2009. IEEE Asian*, 2009, pp. 141–144.

[24] B. Goll and H. Zimmermann, "A comparator with reduced delay time in 65-nm cmos for supply voltages down to 0.65 v," *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 56, no. 11, pp. 810–814, Nov 2009.

[25] M. Miyahara, Y. Ysada, D. Paik, and A. Matsuzawa, "A low-noise self-calibrating dynamic comparator for high-speed ADCs," in *Solid-State Circuits Conference, 2008. A-SSCC 2008. IEEE Asian*, 2008, pp. 269–272.

[26] D. Paik, M. Miyahara, and A. Matsuzawa, "An analysis on a dynamic amplifier and calibration methods for a pseudo-differential dynamic comparator," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E95.A, no. 2, pp. 456–470, 2012.

[27] J. G. Truxal, *Automatic Feedback Control System Synthesis*. New York: McGraw-Hill, 1955.

[28] R. C. Jaeger and R. M. Fox, "Phase Plane Analysis of the Upset Characteristics of CMOS RAM Cells," in *Sixth Biennial University/Government/Industry Microelectronics Symp.*, Auburn, AL, 1985, pp. 183–187.

[29] K. Jeppson, "Comments on the metastable behavior of mismatched CMOS latches," *IEEE J. of Solid-State Circuits*, vol. 31, no. 2, pp. 275–277, 1996.

[30] W. A. M. Van Noije, W. Liu, and J. Navarro, S.J., "Precise final state determination of mismatched CMOS latches," *IEEE J. of Solid-State Circuits*, vol. 30, no. 5, pp. 607–611, 1995.

[31] R. Middlebrook, *Differential Amplifiers*. New York: Wiley, 1963.

[32] E. A. Vittoz, "Dynamic Analog Techniques," in *Design of VLSI Circuits for Telecommunications and Signal Processing*, 2nd ed., J. E. Franca and Y. P. Tsividis, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1994, ch. 6.

[33] H. Taub and D. L. Schilling, *Principles of Communication Systems*, 2nd ed. New York: McGraw-Hill, 1986.

[34] B. Razavi, "The strongarm latch [a circuit for all seasons]," *IEEE Solid-State Circuits Magazine*, vol. 7, no. 2, pp. 12–17, Spring 2015.

[35] M. Lee and A. A. Abidi, "A 9 b, 1.25 ps Resolution Coarse-Fine Time-to-Digital Converter in 90 nm CMOS that Amplifies a Time Residue," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 4, pp. 769–777, April 2008.

[36] L. Chen, A. Sanyal, J. Ma, X. Tang, and N. Sun, "Comparator common-mode variation effects analysis and its application in SAR ADCs," in *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, Canada, May 2016, pp. 2014–2017.

[37] H. Huang, H. Xu, B. Elies, and Y. Chiu, "A Non-Interleaved 12-b 330-MS/s Pipelined-SAR ADC With PVT-Stabilized Dynamic Amplifier Achieving Sub-1-dB SNDR Variation," *IEEE J. of Solid-State Circuits*, vol. 52, no. 12, pp. 3235–3247, 2017.

[38] P. M. Figueiredo and J. C. Vital, "Kickback noise reduction techniques for CMOS latched comparators," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 53, no. 7, pp. 541–545, July 2006.

[39] M. J. M. Pelgrom, *Analog-to-digital conversion*, 3rd ed. Dordrecht: Springer, 2016.

[40] T. Iizuka and A. A. Abidi, "Comprehensive Analysis of Distortion in the Passive FET Sample-and-Hold Circuit," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 4, pp. 1325–1336, Sept 2017.

[41] M.-J. E. Lee, W. J. Dally, and P. Chiang, "Low-power area-efficient high-speed I/O circuit techniques," *IEEE J. of Solid-State Circuits*, vol. 35, no. 11, pp. 1591–1599, 2000.

[42] L. Chen, X. Tang, A. Sanyal, Y. Yoon, J. Cong, and N. Sun, "A 0.7-V 0.6- μW 100-kS/s Low-Power SAR ADC With Statistical Estimation-Based Noise Reduction," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 5, pp. 1388–1398, May 2017.

[43] J. A. Fredenburg and M. P. Flynn, "A 90-MS/s 11-MHz-Bandwidth 62-dB SNDR Noise-Shaping SAR ADC," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 12, pp. 2898–2904, Dec 2012.

[44] Z. Chen, M. Miyahara, and A. Matsuzawa, "A 9.35-ENOB, 14.8 fJ/conv.-step fully-passive noise-shaping SAR ADC," in *Symp, on VLSI Circuits*, Kyoto, Japan, June 2015, pp. C64–C65.

[45] P. Figueiredo, "Comparator metastability in the presence of noise," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 60, no. 5, pp. 1286–1299, May 2013.

[46] T. Iizuka and A. A. Abidi, "FET-R-C Circuits: A Unified Treatment - Part II: Extension to Multi-Paths, Noise Figure, and Driving-Point Impedance," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 63, no. 9, pp. 1337–1348, Sept 2016.

[47] E. Hegazi and A. A. Abidi, "Varactor characteristics, oscillator tuning curves, and am-fm conversion," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 6, pp. 1033–1039, June 2003.

[48] R. J. Schwarz and B. Friedland, *Linear Systems*. New York: McGraw-Hill, 1965.

[49] A. T. Ramkaj, M. Strackx, M. S. J. Steyaert, and F. Tavernier, "A 1.25-GS/s 7-b SAR ADC With 36.4-dB SNDR at 5GHz Using Switch-Bootstrapping, USPC DAC and Triple-Tail Comparator in 28-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 7, pp. 1889–1901, July 2018.

[50] C. C. Liu, S. J. Chang, G. Y. Huang, and Y. Z. Lin, "A 10-bit 50-MS/s SAR ADC With a Monotonic Capacitor Switching Procedure," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 4, pp. 731–740, April 2010.

[51] B. Verbruggen, M. Iriguchi, and J. Craninckx, "A 1.7 mW 11b 250 MS/s 2-Times Interleaved Fully Dynamic Pipelined SAR ADC in 40 nm Digital CMOS," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 12, pp. 2880–2887, Dec 2012.

[52] H. S. Bindra, C. E. Lokin, D. Schinkel, A. J. Annema, and B. Nauta, "A 1.2-V Dynamic Bias Latch-Type Comparator in 65-nm CMOS With 0.4-mV Input Noise," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 7, pp. 1902–1912, 2018.

[53] T. Sepke, P. Holloway, C. Sodini, and H.-S. Lee, "Noise analysis for comparator-based circuits," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 56, no. 3, pp. 541–553, March 2009.

[54] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill, 1984.

[55] R. Staszewski, J. Wallberg, S. Rezeq, C.-M. Hung, O. Eliezer, S. Vemulapalli, C. Fernando, K. Maggio, R. Staszewski, N. Barton, M.-C. Lee, P. Cruise, M. Entezari, K. Muhammad, and D. Leipold, "All-Digital PLL and Transmitter for Mobile Phones," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2469–2482, Dec 2005.

[56] N. Da Dalt, E. Thaller, P. Gregorius, and L. Gazsi, "A Compact Triple-Band Low-Jitter Digital LC PLL with Programmable Coil in 130-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 40, no. 7, pp. 1482–1490, July 2005.

[57] J. Zhuang, Q. Du, and T. Kwasniewski, "A 4GHz Low Complexity ADPLL-based Frequency Synthesizer in 90nm CMOS," in *IEEE CICC Proc.*, Sept 2007, pp. 543–546.

[58] A. Rylyakov, J. Tierno, G. English, M. Sperling, and D. Friedman, "A Wide Tuning Range (1 GHz-to-15 GHz) Fractional-N All-Digital PLL in 45nm SOI," in *IEEE CICC Proc.*, Sept 2008, pp. 431–434.

[59] M. Lee, M. Heidari, and A. A. Abidi, "A Low-Noise Wideband Digital Phase-Locked Loop Based on a Coarse-Fine Time-to-Digital Converter With Subpicosecond Resolution," *IEEE J. Solid-State Circuits*, vol. 44, no. 10, pp. 2808–2816, Oct 2009.

[60] N. Pavlovic and J. Bergervoet, "A 5.3GHz Digital-to-Time-Converter-Based Fractional-N All-Digital PLL," in *IEEE ISSCC Dig.*, Feb 2011, pp. 54–56.

[61] L. Vercesi, L. Fanori, F. De Bernardinis, A. Liscidini, and R. Castello, "A Dither-Less All Digital PLL for Cellular Transmitters," *IEEE J. Solid-State Circuits*, vol. 47, no. 8, pp. 1908–1920, Aug 2012.

[62] Z. Ru, P. Geraedts, E. Klumperink, X. He, and B. Nauta, "A 12GHz 210fs 6mW Digital PLL with Sub-Sampling Binary Phase Detector and Voltage-Time Modulated DCO," in *IEEE Symp. on VLSI Circuits (VLSIC)*, June 2013, pp. C194–C195.

[63] J. Cessna and D. Levy, "Phase Noise and Transient Times for a Binary Quantized Digital Phase-Locked Loop in White Gaussian Noise," *IEEE Trans. on Communications*, vol. 20, no. 2, pp. 94–104, Apr. 1972.

[64] R. C. Walker, "Designing Bang-bang PLLs for Clock and Data Recovery in Serial Data Transmission Systems," in *Phase-Locking in High-Performance Systems-From Devices to Architectures*, B. Razavi, Ed. IEEE Press, 2003, pp. 34–45.

[65] B. Razavi, *RF Microelectronics*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2012.

[66] M.-J. Park and J. Kim, "Pseudo-Linear Analysis of Bang-Bang Controlled Timing Circuits," *IEEE Trans. Circuits Syst. I*, vol. 60, no. 6, pp. 1381–1394, June 2013.

[67] T. K. Kuan and S. I. Liu, "A Bang Bang Phase-Locked Loop Using Automatic Loop Gain Control and Loop Latency Reduction Techniques," *IEEE J. of Solid-State Circuits*, vol. 51, no. 4, pp. 821–831, April 2016.

[68] N. Da Dalt, "Markov Chains-Based Derivation of the Phase Detector Gain in Bang-Bang PLLs," *IEEE Trans. Circuits Syst. II*, vol. 53, no. 11, pp. 1195–1199, Nov 2006.

[69] J. Holmes, "Performance of a First-Order Transition Sampling Digital Phase-Locked Loop using Random-Walk Models," *IEEE Trans. on Communications*, vol. 20, no. 2, pp. 119–131, Feb. 1972.

[70] P. Westlake, "Digital Phase Control Techniques," *IRE Trans. on Communications Systems*, vol. 8, no. 4, pp. 237–246, Dec. 1960.

[71] J. Lee, K. Kundert, and B. Razavi, "Analysis and Modeling of Bang-Bang Clock and Data Recovery Circuits," *IEEE J. Solid-State Circuits*, vol. 39, no. 9, pp. 1571–1580, Sept 2004.

[72] J. Sonntag and J. Stonick, "A Digital Clock and Data Recovery Architecture for Multi-Gigabit/s Binary Links," *IEEE J. Solid-State Circuits*, vol. 41, no. 8, pp. 1867–1875, Aug 2006.

[73] N. Da Dalt, "A Design-Oriented Study of the Nonlinear Dynamics of Digital Bang-Bang PLLs," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 52, no. 1, pp. 21–31, Jan 2005.

[74] D. Liu, P. Basedau, M. Helfenstein, J. Wei, T. Burger, and Y. Chen, "A Frequency-Based Model for Limit Cycle and Spur Predictions in Bang-Bang All Digital PLL," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 59, no. 6, pp. 1205–1214, June 2012.

[75] N. Da Dalt, "Linearized Analysis of a Digital Bang-Bang PLL and its Validity Limits Applied to Jitter Transfer and Jitter Generation," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 55, no. 11, pp. 3663–3675, Dec 2008.

[76] M. Zanuso, D. Tasca, S. Levantino, A. Donadel, C. Samori, and A. Lacaita, "Noise Analysis and Minimization in Bang-Bang Digital PLLs," *IEEE Trans. Circuits Syst. II: Express Briefs*, vol. 56, no. 11, pp. 835–839, Nov 2009.

[77] G. Marucci, S. Levantino, P. Maffezzoni, and C. Samori, "Analysis and Design of Low-Jitter Digital Bang-Bang Phase-Locked Loops," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 61, no. 1, pp. 26–36, Jan 2014.

[78] R. C. Booton, Jr., "Nonlinear Control Systems with Random Inputs," *IRE Trans. on Circuit Theory*, vol. 1, no. 1, pp. 9–18, 1954.

[79] R. Price, "A Useful Theorem for Nonlinear Devices Having Gaussian Inputs," *IRE Trans. Information Theory*, vol. 4, no. 2, pp. 69–72, June 1958.

[80] H. Rowe, "Memoryless Nonlinearities with Gaussian Inputs: Elementary Results," *Bell System Tech. J.*, vol. 61, no. 7, pp. 1519–1525, Sept 1982.

[81] S. J. Mason and H. J. Zimmermann, *Electronic circuits, signals, and systems.* New York: Wiley, 1960.

[82] J. S. Bendat and A. G. Piersol, *Engineering applications of correlation and spectral analysis*, 2nd ed. New York: Wiley, 1993.

[83] H. Taub and D. L. Schilling, *Principles of communication systems*, 2nd ed. New York: McGraw-Hill, 1971.

[84] S. Tertinek, J. P. Gleeson, and O. Feely, "Binary phase detector gain in bang-bang phase-locked loops with DCO jitter," *IEEE Trans. on Circuits and Systems II*, vol. 57, no. 12, pp. 941–945, 2010.

[85] J. H. Fischer, "Noise Sources and Calculation Techniques for Switched Capacitor Filters," *IEEE J. of Solid-State Circuits*, vol. SC-17, no. 4, pp. 742–752, Aug. 1982.

[86] M. H. Perrott, M. D. Trott, and C. G. Sodini, "A modeling approach for Σ-Δ fractional-N frequency synthesizers allowing straightforward noise analysis," *IEEE J. of Solid-State Circuits*, vol. 37, no. 8, pp. 1028–1038, 2002.

[87] R. Nonis, W. Grollitsch, T. Santa, D. Cherniak, and N. D. Dalt, "digPLL-Lite: A Low-Complexity, Low-Jitter Fractional-N Digital PLL Architecture," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 12, pp. 3134–3145, Dec 2013.

[88] J. K. Roberge, *Operational Amplifiers: Theory and Practice.* New York: Wiley, 1975.

[89] A. J. Viterbi, *Principles of Coherent Communications.* New York: McGraw-Hill, 1966.

[90] A. A. Abidi, "Phase Noise and Jitter in CMOS Ring Oscillators," *IEEE J. of Solid-State Circuits*, vol. 41, no. 8, pp. 1803–1816, Aug 2006.

[91] G. R. MacLane, "Table of Integrals," in *Theory of Servomechanisms*, ser. MIT Radiation Laboratory, H. M. James, N. M. Nichols, and R. S. Phillips, Eds. New York, NY: McGraw-Hill, 1947, vol. 25, pp. 369–370.

[92] S. Pamarti, L. Jansson, and I. Galton, "A wideband 2.4-GHz delta-sigma fractional-NPLL with 1-Mb/s in-loop modulation," *IEEE J. of Solid-State Circuits*, vol. 39, no. 1, pp. 49–62, Jan 2004.

[93] X. Gao, E. A. M. Klumperink, M. Bohsali, and B. Nauta, "A Low Noise Sub-Sampling PLL in Which Divider Noise is Eliminated and PD/CP Noise is Not Multiplied by N^2 ," *IEEE J. of Solid-State Circuits*, vol. 44, no. 12, pp. 3253–3263, Dec 2009.

[94] S. Jang, S. Kim, S. Chu, G. Jeong, Y. Kim, and D. Jeong, "An All-Digital Bang-Bang PLL Using Two-Point Modulation and Background Gain Calibration for Spread Spectrum Clock Generation," in *IEEE Symp. on VLSI Circuits (VLSIC)*, June 2015, pp. C138–C139.

[95] F. M. Gardner, "Charge-Pump Phase-Lock Loops," *IEEE Trans. Comm.*, vol. 28, no. 11, pp. 1849–1858, Nov 1980.

[96] J. Hein and J. Scott, "Z-Domain Model for Discrete-Time PLL's," *IEEE Trans. Circuits Syst.*, vol. 35, no. 11, pp. 1393–1400, Nov 1988.

[97] S. Vamvakos, V. Stojanovic, and B. Nikolic, "Discrete-Time, Linear Periodically Time-Variant Phase-Locked Loop Model for Jitter Analysis," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 58, no. 6, pp. 1211–1224, June 2011.

[98] R. Staszewski, D. Leipold, C.-M. Hung, and P. Balsara, "TDC-Based Frequency Synthesizer for Wireless Applications," in *IEEE RFIC Symp.*, June 2004, pp. 215–218.

[99] R. B. Staszewski, C. Fernando, and P. T. Balsara, "Event-driven Simulation and modeling of phase noise of an RF oscillator," *IEEE Trans. on Circuits and Systems I*, vol. 52, no. 4, pp. 723–733, April 2005.

[100] R. B. Staszewski, K. Muhammad, D. Leipold, C.-M. Hung, Y.-C. Ho, J. L. Wallberg, C. Fernando, K. Maggio, R. Staszewski, T. Jung, J. Koh, S. John, I. Y. Deng, V. Sarda, O. Moreira-Tamayo, V. Mayega, R. Katz, O. Friedman, O. E. Eliezer, E. de Obaldia, and P. T. Balsara, "All-digital TX frequency synthesizer and discrete-time receiver for Bluetooth radio in 130-nm CMOS," *IEEE J. of Solid-State Circuits*, vol. 39, no. 12, pp. 2278–2291, Dec 2004.

[101] E. Hegazi and A. A. Abidi, "A 17-mW transmitter and frequency synthesizer for 900-MHz GSM fully integrated in 0.35-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 38, no. 5, pp. 782–792, May 2003.

[102] S. Levantino, G. Marzin, and C. Samori, "An Adaptive Pre-Distortion Technique to Mitigate the DTC Nonlinearity in Digital PLLs," *IEEE J. of Solid-State Circuits*, vol. 49, no. 8, pp. 1762–1772, Aug 2014.

[103] B. D. Sahoo and B. Razavi, "A 10-b 1-GHz 33-mW CMOS ADC," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 6, pp. 1442–1452, June 2013.

[104] J. Wu, A. Chou, C. H. Yang, Y. Ding, Y. J. Ko, S. T. Lin, W. Liu, C. M. Hsiao, M. H. Hsieh, C. C. Huang, J. J. Hung, K. Y. Kim, M. Le, T. Li, W. T. Shih, A. Shrivastava, Y. C. Yang, C. Y. Chen, and H. S. Huang, "A 5.4GS/s 12b 500mW pipeline ADC in 28nm CMOS," in *2013 Symposium on VLSI Circuits*, June 2013, pp. C92–C93.

[105] J. Wu, C. Y. Chen, T. Li, L. He, W. Liu, W. T. Shih, S. S. Tsai, B. Chen, C. S. Huang, B. J. J. Hung, H. T. Hung, S. Jaffe, L. K. Tan, and H. Vu, "A 240-mW 2.1-GS/s 52-dB SNDR Pipeline ADC Using MDAC Equalization," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 8, pp. 1818–1828, Aug 2013.

[106] J. Cao, B. Zhang, U. Singh, D. Cui, A. Vasani, A. Garg, W. Zhang, N. Kocaman, D. Pi, B. Raghavan, H. Pan, I. Fujimori, and A. Momtaz, "A 500 mW ADC-Based CMOS AFE With Digital Calibration for 10 Gb/s Serial Links Over KR-Backplane and Multimode Fiber," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 6, pp. 1172–1185, June 2010.

[107] V. H. C. Chen and L. Pileggi, "A 69.5 mW 20 GS/s 6b Time-Interleaved ADC With Embedded Time-to-Digital Calibration in 32 nm CMOS SOI," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 12, pp. 2891–2901, Dec 2014.

[108] D. Cui, H. Zhang, N. Huang, A. Nazemi, B. Catli, H. G. Rhew, B. Zhang, A. Momtaz, and J. Cao, "3.2 A 320mW 32Gb/s 8b ADC-based PAM-4 analog front-end with programmable gain control and analog peaking in 28nm CMOS," in *2016 IEEE International Solid-State Circuits Conference (ISSCC)*, Jan 2016, pp. 58–59.

[109] J. Cao, D. Cui, A. Nazemi, T. He, G. Li, B. Catli, M. Khanpour, K. Hu, T. Ali, H. Zhang, H. Yu, B. Rhew, S. Sheng, Y. Shim, B. Zhang, and A. Momtaz, "29.2 A transmitter and receiver for 100Gb/s coherent networks with integrated 4x64GS/s 8b ADCs and DACs in 20nm CMOS," in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, Feb 2017, pp. 484–485.

[110] C. H. Chan, Y. Zhu, S. W. Sin, U. Seng-Pan, R. P. Martins, and F. Maloberti, "A 7.8-mW 5-b 5-GS/s Dual-Edges-Triggered Time-Based Flash ADC," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 8, pp. 1966–1976, Aug 2017.

[111] H. Huang, L. Du, and Y. Chiu, "A 1.2-GS/s 8-bit Two-Step SAR ADC in 65-nm CMOS With Passive Residue Transfer," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 6, pp. 1551–1562, June 2017.

[112] B. Xu, Y. Zhou, and Y. Chiu, "A 23-mW 24-GS/s 6-bit Voltage-Time Hybrid Time-Interleaved ADC in 28-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 4, pp. 1091–1100, April 2017.

[113] C. H. Chan, Y. Zhu, I. M. Ho, W. H. Zhang, S. P. U, and R. P. Martins, "16.4 A 5mW 7b 2.4GS/s 1-then-2b/cycle SAR ADC with background offset calibration," in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*, Feb 2017, pp. 282–283.

[114] D. Stepanovic and B. Nikolic, "A 2.8 GS/s 44.6 mW Time-Interleaved ADC Achieving 50.9 dB SNDR and 3 dB Effective Resolution Bandwidth of 1.5 GHz in 65 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 4, pp. 971–982, April 2013.

[115] M. Brandolini, Y. J. Shin, K. Raviprakash, T. Wang, R. Wu, H. M. Geddada, Y. J. Ko, Y. Ding, C. S. Huang, W. T. Shih, M. H. Hsieh, A. W. T. Chou, T. Li, A. Shrivastava, D. Y. C. Chen, B. J. J. Hung, G. Cusmai, J. Wu, M. M. Zhang, Y. Yao, G. Unruh, A. Venes, H. S. Huang, and C. Y. Chen, "A 5 GS/s 150 mW 10 b SHA-Less Pipelined/SAR Hybrid ADC for Direct-Sampling Systems in 28 nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 12, pp. 2922–2934, Dec 2015.

[116] S. Lee, A. P. Chandrakasan, and H. S. Lee, "A 1 GS/s 10b 18.9 mW Time-Interleaved SAR ADC With Background Timing Skew Calibration," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 12, pp. 2846–2856, Dec 2014.

[117] D. J. Huber, "A 10b 160Mb/s 84mW Subranging ADC in 1V 90nm CMOS," Master's thesis, University of California, Los Angeles, 2008.

[118] R. J. Chandler, "Design Techniques for High-Speed, Low-Area Subranging Anlog-to-Digital Converters in 90nm CMOS," Master's thesis, University of California, Los Angeles, 2008.

[119] C. Enz, F. Chicco, and A. Pezzotta, "Nanoscale mosfet modeling: Part 1: The simplified ekv model for the design of low-power analog circuits," *IEEE Solid-State Circuits Magazine*, vol. 9, no. 3, pp. 26–35, Summer 2017.

[120] H. K. Hong, H. W. Kang, B. Sung, C. H. Lee, M. Choi, H. J. Park, and S. T. Ryu, "An 8.6 enob 900ms/s time-interleaved 2b/cycle sar adc with a 1b/cycle reconfiguration for resolution enhancement," in *2013 IEEE International Solid-State Circuits Conference Digest of Technical Papers*, Feb 2013, pp. 470–471.

[121] H. K. Hong, W. Kim, H. W. Kang, S. J. Park, M. Choi, H. J. Park, and S. T. Ryu, "A Decision-Error-Tolerant 45 nm CMOS 7b 1 GS/s Nonbinary 2b/Cycle SAR ADC," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 2, pp. 543–555, Feb 2015.

[122] M. El-Chammas and B. Murmann, "General analysis on the impact of phase-skew in time-interleaved adcs," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 5, pp. 902–910, May 2009.