DYNAMIC AMPLIFIERS FOR HIGH-SPEED PIPELINED A/D CONVERSION

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Ray Nguyen

December 2011

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

Boris Murmann (Principal Advisor)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

Bruce A. Wooley

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

_____

Mark Horowitz

Approved for the University Committee on Graduate Studies

ABSTRACT

Analog-to-digital converters (ADC) are a vital part of a many applications that require an interface with real-world analog signals. Fueled by the ever increasing demand for higher bandwidth and lower power consumption in many areas, the energy efficiency of ADCs becomes a critical performance criterion. Today, there exist a variety of ADCs that provide high energy efficiency solutions only for low bandwidths (below ~100 MHz). In the high-speed space (above 100 MHz), however, the energy efficiency of ADCs degrades dramatically, and this especially visible for pipelined ADCs, which take 3-5 times more energy than other architectures that do not emphasize high speed. Furthermore, existing non-pipelined solutions for this bandwidth range are few in numbers, and this presents an opportunity for innovation at both the architectural and circuit design level.

This thesis explores a pipelined ADC design that employs a variety of low-power techniques such as dynamic residue amplification and incomplete settling in a unique way to maximize the speed while maintaining low energy (98 fJ/conv-step). The resulting work advances the state-of-the-art by simultaneously achieving a high conversion rate (500 MS/s), low power (5.1 mW), moderate resolution (8 bits), and low input capacitance (55 fF). The experimental converter was implemented in a 65-nm Silicon-on-Insulator (SOI) CMOS process and is among the first high-performance ADCs employing this technology.

ACKNOWLEDGEMENTS

My time at Stanford has been one of the best and most enjoyable journeys of my life. I have grown both professionally and socially, thanks to the experience that I had.

First and foremost, it is my pleasure to express my deepest gratitude for my advisor, Professor Boris Murmann, for everything that he has taught me and done for me over the years. His patience and knowledge have not only helped me grow tremendously professionally but also expanded my horizon and understanding of what it means to be a great researcher. I cannot begin to imagine how the invaluable skills that I have learnt from you will help me find my future paths and grow beyond the man that I am today.

I also would like to thank Professor Bruce Wooley and Professor Mark Horowitz for being on my dissertation reading committee, and Professor Yoshio Nishi for being on my orals committee. Though briefly, it has been an honor to have interacted and worked with you in the past and to have had you during my PhD qualification examination.

I would really like to express my gratitude to Andreia Cathelin and Christine Raynaud, for all the technical help especially during the hardest time of this project, and STMicroelectronics for chip fabrication.

Also, this project would not have come this far without the help and technical discussion with Ian Young and Sopan Joshi. You both have helped us greatly during the early days of our project.

For all of us who has ever been a part of the circuit family at Stanford, we could never forget Ann Guerra, for her great dedication to aid us in every way possible, and Joseph Little, for keeping our projects “alive” despite the numerous server failures that we experienced. You both were truly the life savers of this family.

Next, I would like to thank my colleagues from the Murmann group for all the social interaction and technical discussions that have helped me go this far. This includes Noam Dolev, Drew Hall, Clay Daigle, Alirezah Dastgheib, Jason Hu, Donghyun Kim, Manar El-Chammas, Pedram Lajevardi, Martin Kramer, Alex Guo, Ross Walker, Vaibhav Tripathi, Siddarth Seth, Bill Chen, Ryan Boesch, Douglas Adams, Man-Chia Chen, Jonathon Spaulding, Mahmood Saadat, and Alex Omid-Zohoor. Many of you have helped me through some of the toughest times at Stanford, and these memories would undoubtedly stay with me wherever I go.

My journey at Stanford would not have been so fulfilling and complete without my great friends. For that, I would like to thank Jayson Jia, Jenny Kwan, Hart Montgomery, Dominic Coey, and Rebecca Siegel for being my dearest weekend buddies, and thank Trang Le, Julianna Prieto, Karen Tran, Liz Wang, Mangala Gopal, and Becca Loew for always keeping me cheerful and true to myself.

To Amie Pham, I would like to thank you for being there for me during one of the most important times of my life, for the great memories, and for your genuine care during the writing of this thesis.

To Kevin and Julie, I would forever be grateful for all the help that you have given me during the past ten years. There are no better cousins that you two.

To my parents and sister, I am eternally grateful for your never ending love and support. I am sorry to have disappointed you for not becoming the medical doctor who saves lives like father; but I hope that this small accomplishment could partly express my gratitude for the countless sacrifices that you made for me. This work and my humble accomplishment today are entirely dedicated to you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1. MOTIVATION

In an age where information and media streaming are readily available at our finger tips, the demand for high media bandwidth on ever "cheaper" battery powered electronics such as mobile smart phones never ceases to increase. From an engineering perspective, such demand translates to a more cost-effective communication system that can deliver the required bandwidth at the lowest possible energy consumption. Since our world is analog by nature, the performance of almost any communication system that interfaces with the real world strongly depends on the efficiency and speed at which analog information can be converted into their digital representations (0's and 1's), so that they can be stored and processed efficiently using digital computers. Responsible for such conversions in not just many critical communication systems but also a great variety of other applications as depicted in Figure 1.1 [1], the analog-to-digital converters (ADC) are indispensable building blocks whose speed and power consumption often limit the performance of the system as a whole.

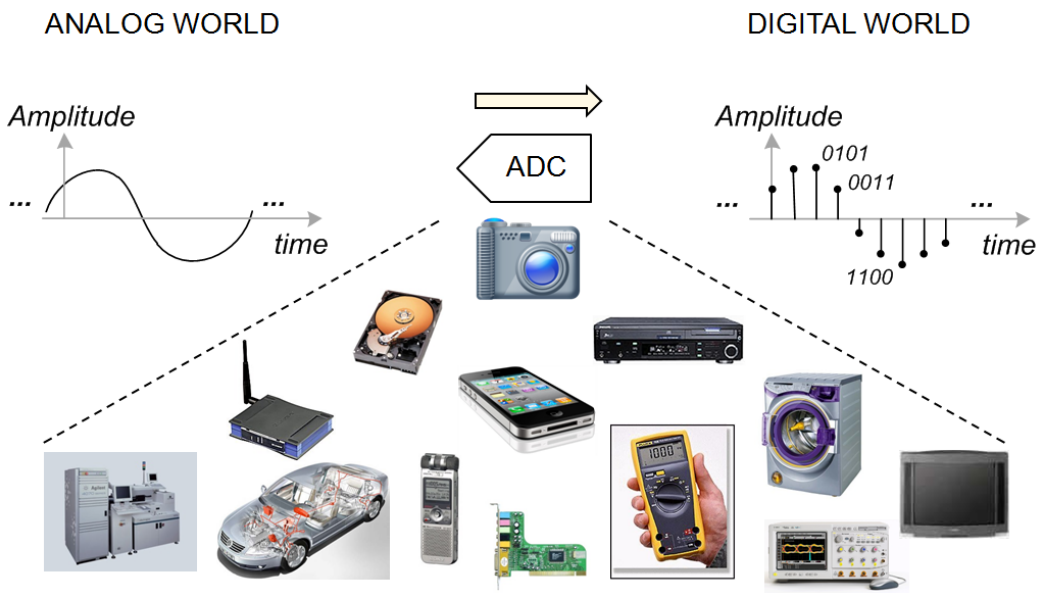


Figure 1.1 - Example applications of ADCs

In addressing this performance issue, tremendous progress has been made in lowering the energy cost of ADCs over the past 14 years (1997-2011). In [2], it was found through a literature survey that the conversion energy of ADCs has on average halved every two years (1997-2009), thanks to a combination of factors such as technology scaling, improved circuit design techniques, and innovative ADC architectures. Today, the choice of ADC architecture for a given application depends largely on the trade-off between its performance and energy cost, with the latter becoming increasingly more critical. To effectively evaluate the performance of the various ADC architectures published over the years, we construct Figure 1.2 using two metrics: the ADC's conversion frequency (f_s) and its energy efficiency, which is defined via a widely used figure-of-merit (FOM) as follows [3]

$$FOM = \frac{P}{2^{ENOB} \cdot f_s} \tag{1.1}$$

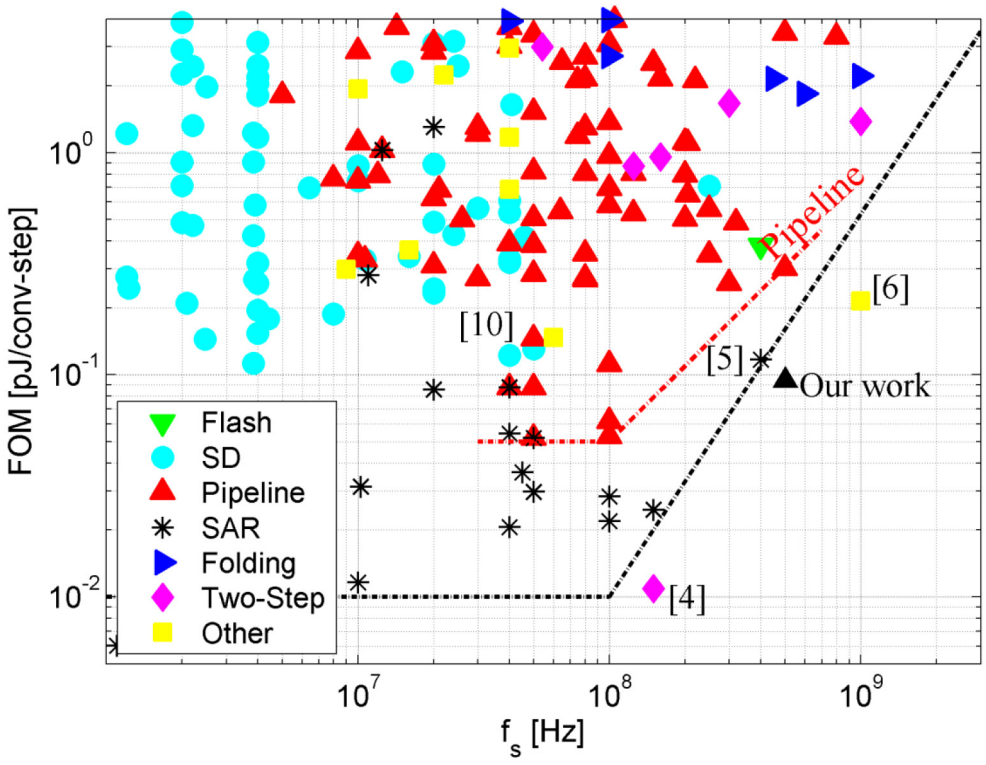


Figure 1.2 – A survey of Published ADCs between 1997 and 2011 (using data from [4])

In this definition of energy efficiency, P represents the power consumption of the ADC, *ENOB* is the effective number of bits which is a (linear) function of the ADC's signal-to-noise-and-distortion ratio, and f_s is the ADC's sample or conversion rate. The unit of FOM is energy per conversion per effective quantization step (J/conv-step). With this definition, a more energy efficient ADC is one with a lower FOM. Although one can argue that this metric is not suitable for very high resolution ADCs (>12 bits) where the cost of increasing one bit of resolution outweighs that of doubling the sampling speed [2], it provides an excellent first-order evaluation of performance for ADCs with low-to-moderate resolutions. To this end and limiting our discussion to non-interleaved architectures with ENOB > 5 bits, Figure 1.2 plots the reported FOM versus the conversion rate f_s on a log-log scale of published data between 1997 and 2011.

By referring to Figure 1.2, we draw two important observations. First, the most energy efficient ADCs published to date are able to achieve an efficiency as low as a few tens of femto-joules (per conversion-step) for f_s below ~150 MS/s. Above this sampling rate, however, it is clear from the survey that the overall energy efficiency degrades dramatically as f_s increases (resembling a familiar "pareto optimal curve" in digital design where investing more energy only provides an increasingly diminishing return in performance [5-6]). Above f_s ~ 150 MS/s, [7], [8], and [9] together represent the state-of-the-art designs at the time of this writing. Compared to these designs, the energy efficiency of pipelined ADCs, despite being a popular architecture for a wide frequency range (5 MHz – GHz range), is about a factor of 3…5 worse. To improve the efficiency of pipelined ADC at high conversion rates, therefore, is the central focus of this research.

1.2. Research Contribution and Thesis Organization

In this thesis, we will present the analysis and design of a high-speed energy efficient 8-bit pipelined ADC implemented in a 65-nm Silicon-on-Insulator (SOI) CMOS technology that advances the state-of-the-art [10] (see Figure 1.2) by simultaneously achieving a high conversion rate (500 MS/s), low power consumption (5.1 mW), and low input capacitance (55 fF). This combination of performance is accomplished via the use of a charge based low-power dynamic amplification technique that was first introduced in 2008 for a 50 MS/s converter [11]. Though the conversion energy of this converter was competitive (FOM = 119 fJ/conv-step), its modest conversion rate of 50 MS/s made the converter's overall performance pale in comparison with other more efficient ADC architectures in this frequency range (e.g. successive approximation converters – SAR), as evidently shown in Figure 1.2. In this work, however, this idea is re-discovered through new analytical insights and re-engineered with additional design techniques to provide a tenfold increase in conversion speed while maintaining low conversion energy. These techniques include the use of incomplete settling [12] and a comparator look-ahead scheme that eliminates the timing overhead of the sub-converters within a pipelined stage. To the

best of our knowledge, this combination of ideas represents a unique and original contribution of our research. As evidently shown in Figure 1.2, our achieved FOM of 98 fJ/conv-step places this design as the most energy efficient converter above 150 MS/s and is a factor of three better than the next best pipelined design within the same range of conversion frequency.

This thesis is organized as follows. Chapter 2 briefly reviews the fundamental concept of pipelined ADCs and residue amplifiers and introduces the basic idea of dynamic amplification. Chapter 3 then delves into the discussion of the dynamic amplifier's design trade-offs. Next, Chapter 4 examines the transient settling of dynamic amplifiers and introduces incomplete settling as a technique to maximize the amplifier's speed. Architectural implementation and the comparator look-ahead technique are subsequently discussed in Chapter 5, followed by the measurement results of our prototype ADC in Chapter 6. Finally, Chapter 7 suggests future work directions and closes our discussion with a conclusion.

CHAPTER 2. RESIDUE AMPLIFIERS IN PIPELINED A/D CONVERSION

Do you need this heading?

2.1. REVIEW OF BASIC OPERATION AND ARCHITECTURE

This section serves as a brief review of pipelined ADCs to establish the groundwork for the discussion that follows. More extensive treatments on the subject can be found elsewhere [13], [14].

A pipelined ADC converts an analog signal V_{in} into its digital representation D_{out} by breaking the conversion task into multiple steps that are handled sequentially by a cascade of ADCs of lower resolution. As shown in Figure 2.1, the structure of each of these coarse ADC stages in the pipeline consists of three primary blocks: a sub ADC, a digital-to-analog converter (DAC), and a residue amplifier with gain G. In the design of pipelined ADCs, a sample-and-hold (SHA) block can optionally be added at the front of the ADC to improve its dynamic performance at high conversion rates [13].

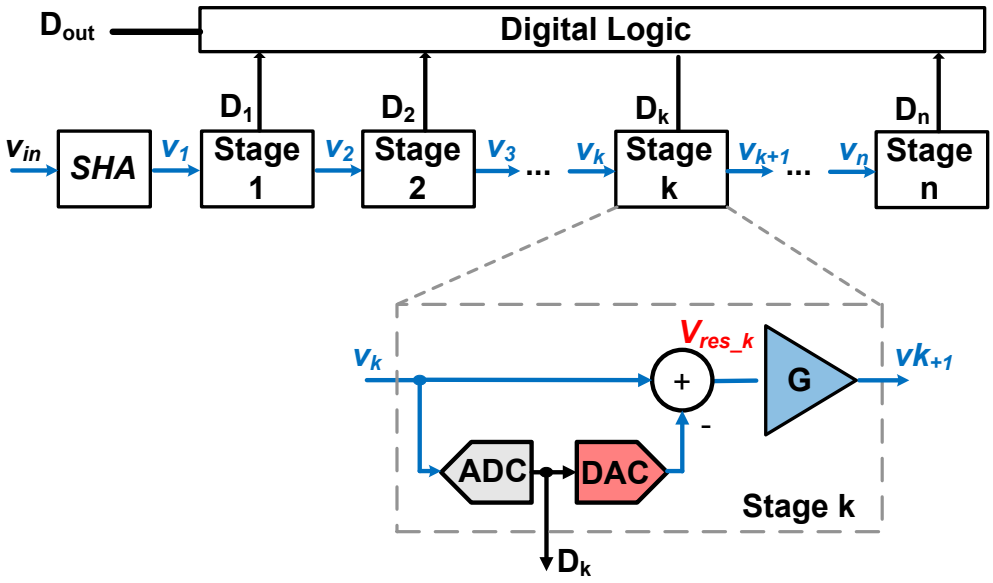


Figure 2.1 – Structure of a pipelined ADC

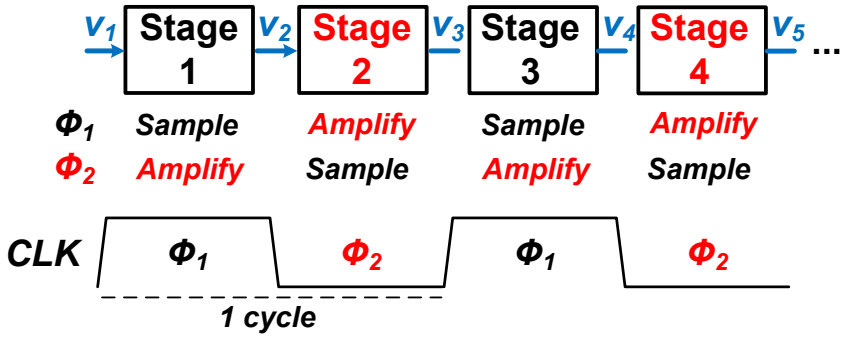


Figure 2.2 - Pipelined operation

The functionality of each stage is as follows. The sub ADC digitizes the input signal v_k to produce the stage's coarse conversion result D_k, which is then returned to the analog domain by the DAC before getting subtracted from the input signal v_k. The result of this difference is the stage's conversion error or residue, V_{res_k}. V_{res_k} is next amplified by the residue amplifier G before being fed as the input signal v_{k+1} into the next stage (i.e. stage k+1) for further processing. Once stage k+1 samples v_{k+1}, the same conversion steps repeat to subsequently produce D_{k+1} and v_{k+2}. This process continues until the last digital output D_n is generated, at which point all of the codes $D_1 \ldots D_n$ can be digitally combined to produce the overall conversion result, D_{out}.

As depicted in Figure 2.2, pipeline ADCs work in a ping pong fashion in which, during the first half ϕ_1 of a clock (CLK) cycle, the even stages 2, 4, 6… amplify their residue signals V_{res} to produce the outputs v_3, v_5, v_7 …, while the following odd stages 3, 5, 7 … respectively sample these same signals at their inputs. During the second half ϕ_2 of the clock cycle, these even stages (2, 4, 6, …) now switch to sampling from the outputs of the preceding odd stages (1, 3, 5, ...) which are now amplifying their residue signals to respectively produce v_2, v_4, v_6... As a new cycle begins, this process repeats. Defining the conversion cycle as the total time for each stage to sample and amplify, this structure thus has a half-cycle latency between the times when D_k is produced by stage k and when D_{k+1} is generated by stage k+1.

As previously demonstrated by the survey in Figure 1.2, although pipelined ADCs are a very popular architecture for applications that demand a moderate-to-high bandwidth (5 MHz – GHz range) and resolution (6-14 bits), their energy efficiency pale in comparison to other architectures (e.g. SAR) that define the state-of-the-art envelope. The cause of this can be attributed to several reasons. The first is due to the high overhead associated with implementing pipelined ADCs. In particular, as seen in Figure 2.1, a complete pipeline stage requires a sub-ADC, sub-DAC, and an amplifier, each of which introduces non-idealities such as offsets, noise, distortion, and delay that occupy a significant portion of the ADC's cycle time and design budget and are costly to optimize. Furthermore, clocking schemes to extract the best possible speed of this type of converter are also complex and often demand high power consumption.

The second and more important reason for the energy inefficiency of a pipelined converter is due to the high power consumption of its stage residue amplifiers. This is evident in the amount of increasingly new research ideas published in the past 8 years that attempt to address the power overhead of residue amplifiers through various techniques. To name a few, some of these techniques include stage scaling [15], amplifier sharing [16], SHA-less front-end [17, 18], comparator based [14–15], open-loop [21–23], class AB [19–20], charge redistribution [7, 21–22] amplification, and incomplete settling [12]. Reducing the amplifier power consumption, therefore, is imperative in improving the energy efficiency of pipelined

ADCs and is also the primary objective of this research. To this end, the focus of the next two sections is a review of the general concept of residue amplifiers.

2.2. CONVENTIONAL RESIDUE AMPLIFIER

The traditional way of amplifying residue signals involves employing a high gain class-A operational amplifier (op amp) in a switched capacitor configuration as shown in Figure 2.3. The residue amplifier operates in 2 phases: sample and amplify. During the sample phase, the input signal V_{in} is sampled onto both capacitors C_1 and C_2, while the input terminals of the op amp are shorted together to an AC ground. The signal dependent charge stored on the negative input terminal of the op amp is thus

$$Q_{sample} = -(C_1 + C_2)V_{in} \tag{2.1}$$

To realize voltage amplification, C_1 is then flipped around and connected across the op amp between its negative input and output, and C_2 is connected to an AC ground as shown in Figure 2.3b. Due to the high loop gain of this feedback network, the negative terminal is forced a virtual ground condition, causing all of the signal dependent charge on C_2 to transfer to C_1. Expressing this mathematically, we have

$$Q_{amplify} = -C_1 V_{out} \tag{2.2}$$

Due to charge conservation on the negative input terminal, Q_{sample} must equal $Q_{amplify}$, leading to expressions of V_{out} and incremental or AC gain given by

$$V_{out} = \frac{C_1 + C_2}{C_1} V_{in} \tag{2.3}$$

$$Gain = \frac{dV_{out}}{dV_{in}} = \frac{C_1 + C_2}{C_1} \tag{2.4}$$

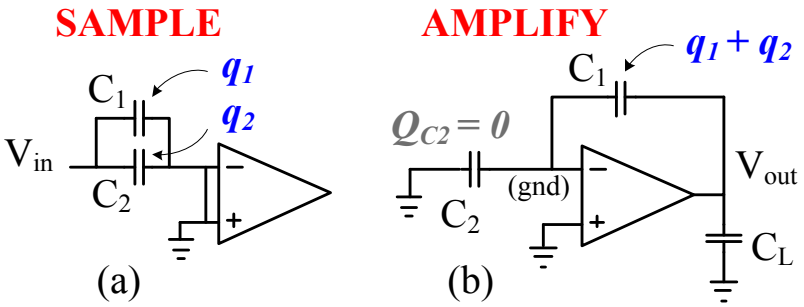


Figure 2.3- Conventional residue amplifier

In short, the traditional residue amplifier relies on the high loop gain of its feedback network to force the sampled charge from C_1 onto C_2 thereby increasing the total signal dependent charge resided on C_2 and causing to the voltage drop across C_2 to be amplified by the ratio given in (2.4). In contrast to this method of voltage amplification, the dynamic amplifier discussed in the next section accomplishes the same charge redistribution without the need for an op amp. As will be apparent in Section 2.4, eliminating the op amps from the pipelined ADCs enables tremendous power saving and significantly improves their energy efficiency.

2.3. Dynamic Residue Amplifier

The principle of dynamic amplification involves employing an NMOS transistor as the main device for both sampling and amplifying residue signals. Figure 2.4 shows the conceptual operation of the amplifier which consists of two phases. During the sampling phase, the NMOS transistor is configured as a MOS capacitor (moscap), and the input signal V_{in} is sampled onto its gate, while its drain and source are tied together and to a fixed voltage V_{bias}. Neglecting the gate-to-body capacitance (C_{gb}) in this first order analysis, the transistor can thus conceptually be broken into C_{gs} and C_{gd} as shown. The sampled charge deposited on the gate in this phase is therefore

$$Q_{sample} = (C_{gs} + C_{gd})(V_{in} - V_{bias}) \tag{2.5}$$

During the amplify phase, the source is connected to a discharged load C_L, and the drain is switched to V_{DD}. The gate is left floating, forcing the transistor to act as a dynamic current source driving C_L, allowing a direct current to flow from V_{DD} through

the transistor and causing V_{out} to rise monotonically. As V_{out} approaches its final value V_{out_f}, the gate-to-source voltage V_{gs} decays monotonically until $V_{gs} \approx V_t$, turning off the current (to first order) and stopping the charging process (see Figure 2.5). At this point, the charge on the gate can be expressed as

$$\begin{aligned} Q_{amplify} &= C_{gs}V_{gs} + C_{gd}(V_g - V_{DD}) \\ &= C_{gd}V_{out} + (C_{gs} + C_{gd})V_t - C_{gd}V_{DD} \end{aligned} \tag{2.6}$$

Due to charge conservation on the floating gate, Q_{sample} must equal $Q_{amplify}$, leading to the following expression for the output voltage

$$V_{out} = \frac{C_{gs} + C_{gd}}{C_{gd}}(V_{in} - V_{bias}) - \frac{C_{gs} + C_{gd}}{C_{gd}}V_t + V_{DD} \tag{2.7}$$

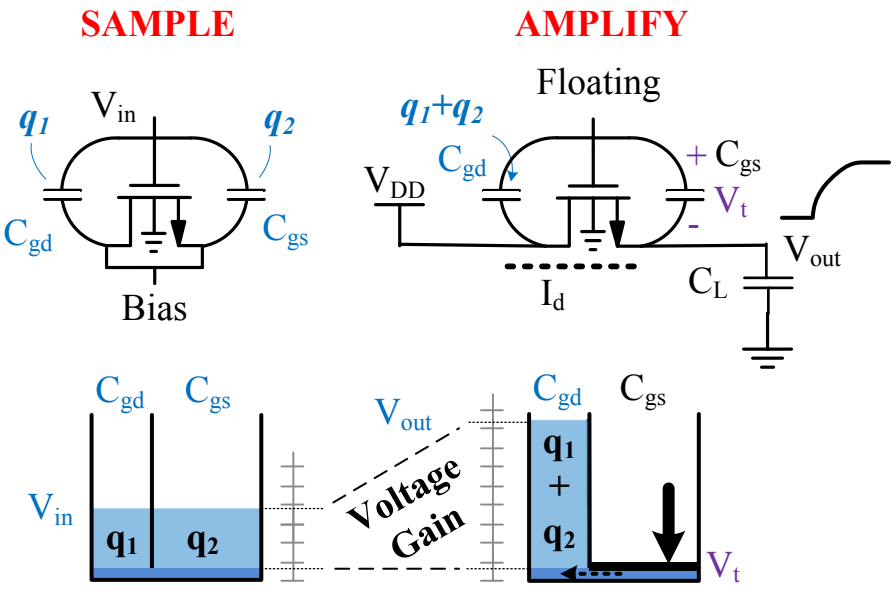


Figure 2.4 - Dynamic Residue Amplifier

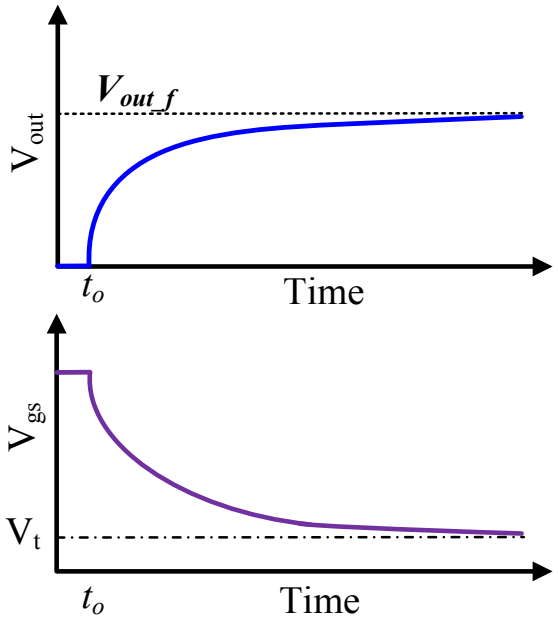


Figure 2.5 - Transient behavior of dynamic amplifier in amplification

Differentiating the above equation leads to an expression for the incremental gain given by

$$Gain = \frac{dV_{out}}{dV_{in}} = \frac{C_{gs} + C_{gd}}{C_{gd}} \tag{2.8}$$

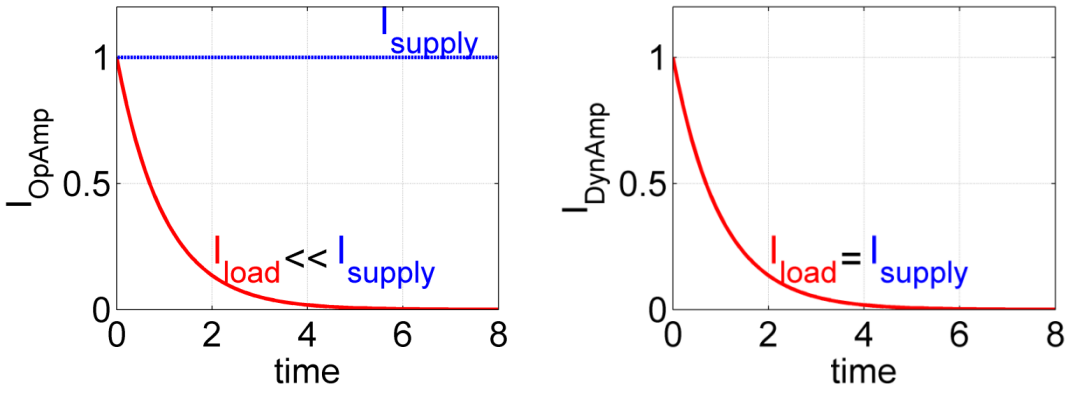
The above result shows that by sampling onto the gate of a moscap and subsequently configuring it as a dynamic current source that conducts current from V_{DD} to a discharged load, incremental voltage amplification can be achieved. Although V_{in} is sampled onto both C_{gs} and C_{gd}, only a constant charge (independent of V_{in}) appears across C_{gs} at the end of amplification. All the signal dependent charge must therefore reside on C_{gd} causing the voltage drop across C_{gd} to increase proportionally. At this point, since the drain is at a DC voltage, and V_{out} is a constant threshold drop below the gate, this incremental AC gain signal also shows up at the output node V_{out}.

To further aid our understanding, a hydraulic model is also included in the bottom half of Figure 2.4. Here, the sampling process can be visualized by viewing charge as a liquid (Q), capacitance as the width of the liquid container (C), and voltage as the liquid level ($V = Q/C$). The voltage amplification is due to pushing a set amount

of liquid (charge proportional to input) from a wide tub (large C) into a smaller tub (smaller C), thereby providing a rise in the liquid level (voltage amplification) in the latter.

2.4. Comparison between Conventional and Dynamic Amplifiers

The greatest advantage in employing dynamic amplifiers for pipelined A/D conversion lies in their power saving benefits. As illustrated in Figure 2.6, a conventional class-A op amp based residue amplifier draws a constant current during its operation, while the current delivered to its load decreases rapidly with time [28]. This leads to low power efficiency. On the contrary, all of the supply current drawn by the dynamic amplifier is delivered to the load. Clearly, to deliver the same amount of charge to an identical load, the dynamic amplifier is far more efficient than the conventional op amp.



(a) (b)

Figure 2.6 – Power efficiency comparison between (a) a class-A OpAmp and (b) a dynamic amplifier

In trading away the many benefits of a feedback amplifier (e.g. high closed-loop bandwidth, suppression of amplifier's nonlinearity, robustness, etc.), however, the design of dynamic amplifiers requires a careful study of the non-idealities associated with such an open-loop class-B topology. Concerns regarding the topology's maximum achievable speed and resolution, robustness, supply rejection, and temperature

dependence undoubtedly arise. Some of these concerns are addressed in the rest of this thesis, while others are deferred for future study. In the next chapter, analytical results are presented to quantify the various tradeoffs associated with this amplifer.

2.5. CHAPTER SUMMARY

This chapter reviewed the basic concept and operation of the popular pipelined ADC architecture. Following this review is a discussion that highlighted the importance of improving the power efficiency of residue amplifiers, the fundamental building blocks in a pipelined ADC that usually dominate its power consumption. The chapter then introduced the concept of dynamic amplification as a much more power efficient alternative to the conventional op amps that are often employed in pipelined ADCs.

CHAPTER 3. ANALYSIS OF DYNAMIC RESIDUE AMPLIFIERS

The previous chapter qualitatively introduced the concept of dynamic amplification to establish some basic intuition on the working of the amplifier. This chapter focuses on the quantitative analysis of the amplifier's tradeoffs in a pipelined ADC. To illustrate and highlight the key observations made throughout this chapter, this work, which uses a 65-nm thin film SOI process, is compared against prior art which was implemented in a 90-nm Bulk CMOS technology [11]. Generally, our approach consists of establishing an intuitive view of the presented tradeoffs using the amplifier's hydraulic model introduced in Chapter 2, followed by quantitative results developed to complement the discussion.

In terms of organization, the chapter begins with a description of the amplifier's differential configuration in Section 3.1, followed by a closer examination of the non-idealities that influence the amplifier's achievable gain in Section 3.2. Section 3.3 then analyzes the amplifier's output common mode and its ability to reject input common mode and supply variations. Following this discussion, Section 3.4 examines the issue

of sampling linearity to highlight the key differences between the Bulk and SOI CMOS amplifiers. Noise is subsequently considered in Section 3.5. Finally, the chapter concludes with a brief summary.

3.1. Pseudo-Differential Implementation

The dynamic amplification concept introduced in Chapter 2 used a single-ended amplifier. In an actual implementation, differential configurations are preferred to minimize the impact of even-order distortions. One can conceptually visualize this using the same hydraulic model of Figure 2.4 but with two sets of tubs to represent a pair of single-ended amplifiers (see Figure 3.1). The input and output signals V_{id} and V_{od} are now taken as the difference in liquid level of the two tubs. As a result, the voltage gain, being the ratio of the differential liquid levels V_{od} and V_{id}, is identical to the incremental gain given by (2.8) and the single-ended hydraulic model of Figure 2.4.

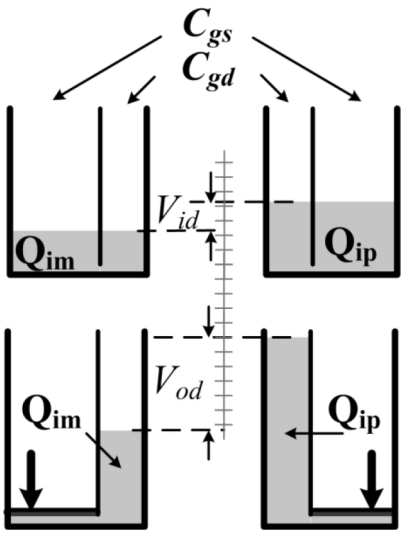


Figure 3.1 – Hydraulic model of an ideal differential dynamic amplifier

To illustrate this more quantitatively, additional circuit analysis is necessary. Specifically, during the sampling phase, the moscaps are connected in a fully differential fashion. The inputs V_{ip} and V_{im} are sampled on the gates, while the drains and sources of both transistors are tied together. Assuming voltage-independent and

perfectly matched moscaps, the equivalent circuit of this differential network can be deduced as shown in Figure 3.2.

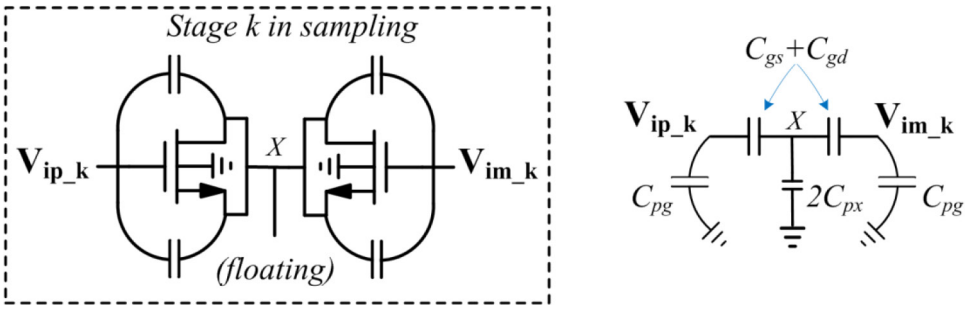


Figure 3.2 - Fully differential sampling structure and its equivalent circuit

In this circuit, node X serves as an AC ground. The parasitic capacitance at this node is labeled as $2C_{px}$ to denote the total parasitic contribution from the drains and sources of both half circuits. Here, C_{pg} is the gate's total parasitic capacitance. The plus and minus charges sampled onto the two gates at the end of the sampling phase are

$$Q_{sample\_ip/im} = (C_{gs} + C_{gd} + C_{pg})V_{ip/im} - (C_{gs} + C_{gd})V_X \qquad (3.1)$$

Next, to implement amplification, these two transistors are reconfigured as separate single-ended amplifiers that drive the next pipelined stage, which is now sampling. The output residue signal is $V_{od}\ (k)$ and is taken as the difference of the two amplifier outputs, $V_{op}\ (k)$ and $V_{om}\ (k)$ as shown in Figure 3.3.

The equivalent differential circuit of Figure 3.3 is the same as what shown in Figure 2.4, in which the main transistor of stage k is in amplification and drives a capacitive load (i.e. stage k+1 in sampling). In this case, the total load seen at its output, as can be seen in Figure 3.2, is $C_L = C_{gs} + C_{gd} + C_{pg}$. Analyzing both circuit halves in the same manner as in Section 2.3 while neglecting the signal dependence of the threshold voltage V_t and transistors' finite output resistances, we get the total plus and minus charges on the gates at the end of amplification as follows

$$Q_{amplify\_p/m} = (C_{pg} + C_{gd})V_{op/om} + (C_{gs} + C_{gd} + C_{pg})V_t - C_{gd}V_{DD} \qquad (3.2)$$

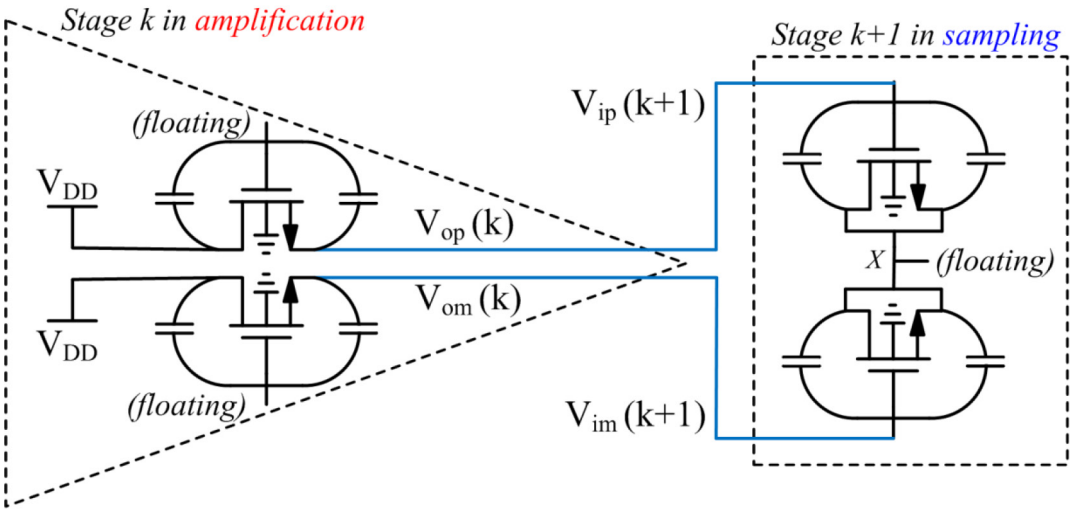


Figure 3.3 - Pseudo-differential amplification followed by fully differential sampling

Due to charge conservation on the gate nodes, Q_{sample} must equal $Q_{amplify}$. One therefore can derive an expression for V_{op} and V_{om} in terms of V_{ip} and V_{im}, similarly to (2.7), to give

$$V_{op/om} = \frac{C_{gs}+C_{gd}+C_{pg}}{C_{gd}+C_{pg}}V_{ip/im} - \frac{C_{gs}+C_{gd}+C_{gp}}{C_{gd}+C_{pg}}V_T - \frac{C_{gs}+C_{gd}}{C_{gd}+C_{pg}}V_X + \frac{C_{gd}}{C_{gd}+C_{pg}}V_{DD} \tag{3.3}$$

The differential mode and common mode outputs of stage k are thus given by

$$V_{od} = \frac{C_{gs}+C_{gd}+C_{pg}}{C_{gd}+C_{pg}}V_{id} \tag{3.4}$$

$$V_{oc} = \frac{C_{pg}+C_{gs}+C_{gd}}{C_{gd}+C_{pg}}V_{ic} - \frac{C_{gs}+C_{gd}}{C_{gd}+C_{pg}}V_X - \frac{C_{gs}+C_{gd}+C_{gp}}{C_{gd}+C_{pg}}V_T + \frac{C_{gd}}{C_{gd}+C_{pg}}V_{DD} \tag{3.5}$$

In line with the prediction from Figure 3.1, (3.4) shows that, in the ideal situation (where perfect matching and signal independent moscaps and threshold voltages are assumed) and in the absence of gate parasitic capacitances ($C_{pg} = 0$), the differential gain ($G = dV_{od}/dV_{id}$) is identical to the incremental gain of the single-ended case given by (2.8). Here, this analysis also took into account gate parasitic

capacitances to show their effects on the amplifier's gain. From the hydraulic point of view of Figure 3.1, any parasitic capacitance that appears on the gates effectively enlarges the C_{gd} tubs. Consequently, for the same liquid volume ΔQ transferred from the C_{gs} tubs to the C_{gd} tubs, this enlargement translates to a proportional reduction in liquid level ΔV in the C_{gd} tubs, decreasing voltage gain as a result.

For a given technology (i.e. fixed C_{gs} and C_{gd}), the gate parasitic capacitance C_{pg} provides an upper bound on the maximum achievable voltage gain of the amplifier. As will be discussed further in the next section, it is crucial to minimize this parameter if the highest possible gain is desired. In terms of the output common mode, (3.5) shows that V_{oc} is a function of not only the input common mode V_{ic} but also the supply voltage V_{DD}, the device's threshold V_t, and the bias voltage of node X. Controlling the output common mode and ensuring that the amplifier has sufficient common mode rejection in a pipelined ADC are the topics of Section 3.3.

3.2. GAIN CONSIDERATIONS

The previous section analyzed the amplifier while neglecting the voltage dependent nature of the moscaps and threshold voltages, and the effect of finite output resistances (r_o). This section examines their impact on the amplifier's gain. Specifically, since V_{ip} and V_{op} differ from V_{im} and V_{om} during sampling and amplification, bias dependent quantities such as C_{gs}, C_{gd}, and V_t are thus different for the two circuit halves. Following the same approach as in the previous section, the following analysis offers both the intuitive and quantitative views of these signal dependent effects.

3.2.1. INTUITIVE VIEW WITH A HYDRAULIC MODEL

We begin by examining the impact of capacitance variation on the amplifier's gain using the differential hydraulic model shown in Figure 3.4. With ideal capacitors (see Figure 3.4a), the proportionalities in the V-Q (voltage to charge) conversion during sampling and Q-V conversion during amplification are constant; i.e. the widths of the tubs do not vary. This results in the voltage gain given by (3.4). In reality,

however, due to the voltage dependence of the MOSCAPs, the tub widths vary with voltage, altering the V-Q and Q-V conversion factors adversely (see Figure 3.4b). For instance, given the bias conditions found in [11], C_{gs} decreases with input voltage, thus resulting in a smaller amount of charge being acquired. Upon redistribution, this leads to a smaller output voltage, and thus reduced voltage gain.

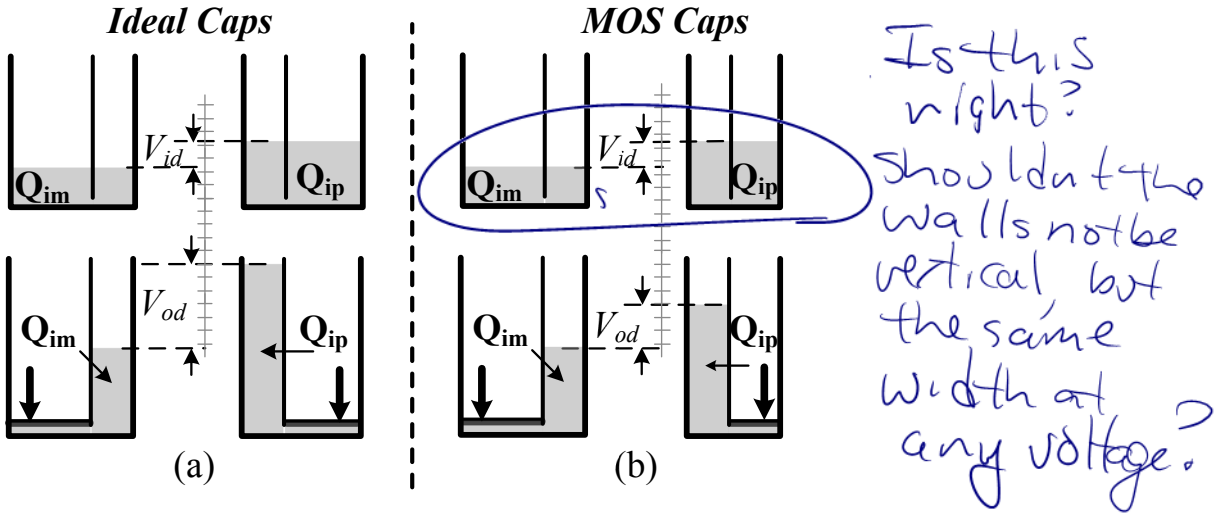


Figure 3.4 - Amplification with (a) Ideal caps and (b) Signal dependent MOSCAPs

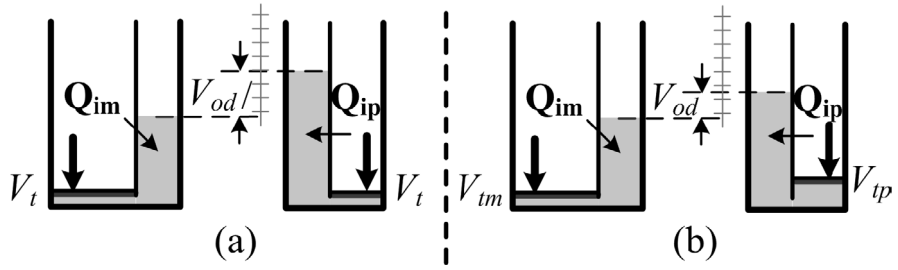


Figure 3.5 - Amplification with (a) Constant threshold voltages and (b) Signal dependent threshold voltages ($V_{tp} > V_{tm}$)

Besides capacitance variation, two other effects also act to worsen the amplifier's gain. Particularly, in the ideal case where the threshold voltages of the two transistors are equal (i.e. signal independent), equal levels of residual liquid (i.e. $V_{gs} \approx V_t$) remain in the two C_{gs} tubs when the amplifiers settle (see Figure 3.5a). Due to the difference in V_{ds} of the two transistors during amplification, both drain induced barrier lowering (DIBL) and finite r_o act to alter the amplifier's current drive strength. In particular, DIBL causes the positive half (one with a higher output $V_{op} > V_{om}$ and hence

larger

a smaller V_{ds}) to always have a V_t than the negative half (i.e. $\Delta V_t = V_{tp} - V_{tm} > 0$). Furthermore, due to finite r_o, a smaller V_{ds} of the positive half requires a larger V_{gs} to maintain the same current, a condition that can be thought of as an additional change in the "effective" threshold voltages that further enlarges ΔV_t, similar to the effect of DIBL. The end result is a higher (lower) liquid level being retained in the positive (negative) tub, as depicted in Figure 3.5b. This difference in liquid level $\Delta V = V_{tp} - V_{tm}$ then becomes a loss in the differential charge ΔQ available for redistribution, further reducing the output signal and hence the amplifier's gain.

3.2.2. Quantitative View with Analytical Results

To complement the hydraulic view in the previous section, quantitative results are presented here. Using the same analysis techniques as in Section 3.1, the plus charge sampled on the positive gate at the end of the sampling phase can be written as

$$Q_{sample\_p} = \left(C^s_{gsp} + C^s_{gdp} + C_{pg}\right)V_{ip} - \left(C^s_{gsp} + C^s_{gdp}\right)V_X \tag{3.6}$$

In the above equation, the superscript s is used to denote that these parameters are from the sampling phase. Using the superscript a to denote the amplification phase, the plus charge on the positive gate at the end of amplification is

$$Q_{amplify\_p} = \left(C_{pg} + C^a_{gdp}\right)V_{op} + \left(C^a_{gsp} + C^a_{gdp} + C_{pgp}\right)V_{tp} - C^a_{gdp}V_{DD} \tag{3.7}$$

Charge conservation again gives us an expression for V_{op} in terms of V_{ip} and other parameters, similarly to (3.3). Here, we make a distinction between the effective positive and negative threshold voltages (i.e. V_{tp} and V_{tm}) to account for their dependence on the transistor biases.

$$V_{op} = \frac{C^s_{gsp} + C^s_{gdp} + C_{pg}}{C_{pg} + C^a_{gdp}}V_{ip} - \frac{C^a_{gsp} + C^a_{gdp} + C_{pg}}{C_{pg} + C^a_{gdp}}V_{tp} - \frac{C^s_{gsp} + C^s_{gdp}}{C_{pg} + C^a_{gdp}}V_X + \frac{C^a_{gdp}}{C_{pg} + C^a_{gdp}}V_{DD} \tag{3.8}$$

$$V_{om} = \frac{C^s_{gsm} + C^s_{gdm} + C_{pg}}{C_{pg} + C^a_{gdm}} V_{ip} - \frac{C^a_{gsm} + C^a_{gdm} + C_{pg}}{C_{pg} + C^a_{gdm}} V_{tm} - \frac{C^s_{gsm} + C^s_{gdm}}{C_{pg} + C^a_{gdm}} V_X + \frac{C^a_{gdm}}{C_{pg} + C^a_{gdm}} V_{DD} \quad (3.9)$$

Taking the difference of V_{op} and V_{om} would yield V_{od}, but such an expression is cumbersome to write due to the various (signal dependent) imbalanced quantities of the positive and negative halves. Instead, let us make a few approximations to get the results that are more in line with the intuition established previously and still stay close to exact analysis. In particular, if we lump C_{gs} and C_{gd} together as a single quantity C_{gsd}, the total gate-to-source and gate-to-drain capacitance, we can then write the following expressions for the positive and negative circuit halves

$$C^s_{gsdp} = C^s_{gsd} + \frac{1}{2}\Delta C^s_{gsd} = C^s_{gsd}\left(1 + \frac{X^s_{gsd}}{2}\right) \quad (3.10)$$

$$C^s_{gsdm} = C^s_{gsd} - \frac{1}{2}\Delta C^s_{gsd} = C^s_{gsd}\left(1 - \frac{X^s_{gsd}}{2}\right) \quad (3.11)$$

$$C^a_{gdp} = C^a_{gd} + \frac{1}{2}\Delta C^a_{gd} = C^a_{gd}\left(1 + \frac{X^a_{gd}}{2}\right) \quad (3.12)$$

$$C^a_{gdm} = C^a_{gd} - \frac{1}{2}\Delta C^a_{gd} = C^a_{gd}\left(1 - \frac{X^a_{gd}}{2}\right) \quad (3.13)$$

$$where \;\; X^s_{gsd} = \frac{\Delta C^s_{gsd}}{C^s_{gsd}} \;\; and \;\; X^a_{gd} = \frac{\Delta C^a_{gd}}{C^a_{gd}}$$

In the above equations, X^s_{gsd} and X^a_{gd} are defined as the normalized or percent variation of C^s_{gsd} and C^a_{gd}, which can be positive or negative. Appendix A shows that, by substituting these expressions into equations (3.8) and (3.9) and taking the difference, the differential output V_{od} can be written as shown in (3.14).

$$V_{od} = G_{ideal}\left[V_{id} - k_c\Delta V_t + \left(X^s_{gsd} - X^a_{gd}\right)\left(V_{ic} - k_c V_{t\_avg} - \frac{C^s_{gsd}}{C^s_{gsd} + C_{pg}} V_X\right)\right] \quad (3.14)$$

$$where\ G_{ideal} = \frac{C_{gsd}^{s}+C_{pg}}{C_{pg}+C_{gd}^{a}},\ \Delta V_t = V_{tp} - V_{tm}, and\ k_c = \frac{C_{gsd}^{a}+C_{pg}}{C_{gsd}^{s}+C_{pg}} (\approx 1 \ldots 2)$$

In the above equation, G_{ideal} represents the familiar first-order gain previously given by (3.4). In the absence of signal dependent imbalances, (3.14) reduces to (3.4). As can be seen, the two major sources of signal dependent imbalance that alter the amplifier's output signal are the threshold imbalance (ΔV_t) and the capacitance imbalance ($X_{gd}^{a} - X_{gsd}^{s}$). Depending on their polarities, they can either add to or subtract from the output signal, effectively altering the amplifier's gain. As previously depicted in Figure 3.5, DIBL and finite output resistances cause V_{id} and ΔV_t to always be of the same polarity, implying that threshold imbalance reduces gain. Capacitor imbalance $X_{gd}^{a} - X_{gsd}^{s}$, on the other hand, can increase or reduce the output signal, depending on the bias voltages and the C-V profile of the moscaps. In [11], this term is negative, leading to a significant gain reduction as previously illustrated in Figure 3.4.

3.2.3. A Numerical Example on the Achievable Gain in Practice

To further illustrate the effects of these non-idealities on the achievable gain, we utilize nominal parameters of the 90-nm CMOS design and assume that $C_{gs} \approx C_{gd}$, $V_{id\_peak} = V_{ic} = \frac{1}{2}V_{dd} = 0.6$ V, $V_x = 0.5$ V, and $V_{t\_avg} = 0.25$ V. Using these values, Figure 3.6 plots the amplifier's effective gain (i.e. $G = V_{od}/V_{id}$ obtained from (3.14) which is less than $G_{ideal} = dV_{od}/dV_{id}$) versus C_{pg} (normalized to C_{gd}) for various imbalance cases of ΔV_t and $X = X_{gsd}^{s} - X_{gd}^{a}$.

In Figure 3.6, it is apparent that the amplifier's gain is sensitive to the gate's parasitic capacitance. In particular, even in the absence of threshold and capacitor imbalances, the bold solid curve shows that a small gate's parasitic capacitance on the order of $C_{pg} \approx C_{gd}$ degrades the maximum achievable G_{ideal} from 2 (where $C_{pg} = 0$) to 1.5 (i.e. a 25 % reduction). This suggests that the circuitry interfacing with the amplifier's gates should be kept to a minimum to mitigate their parasitic contribution.

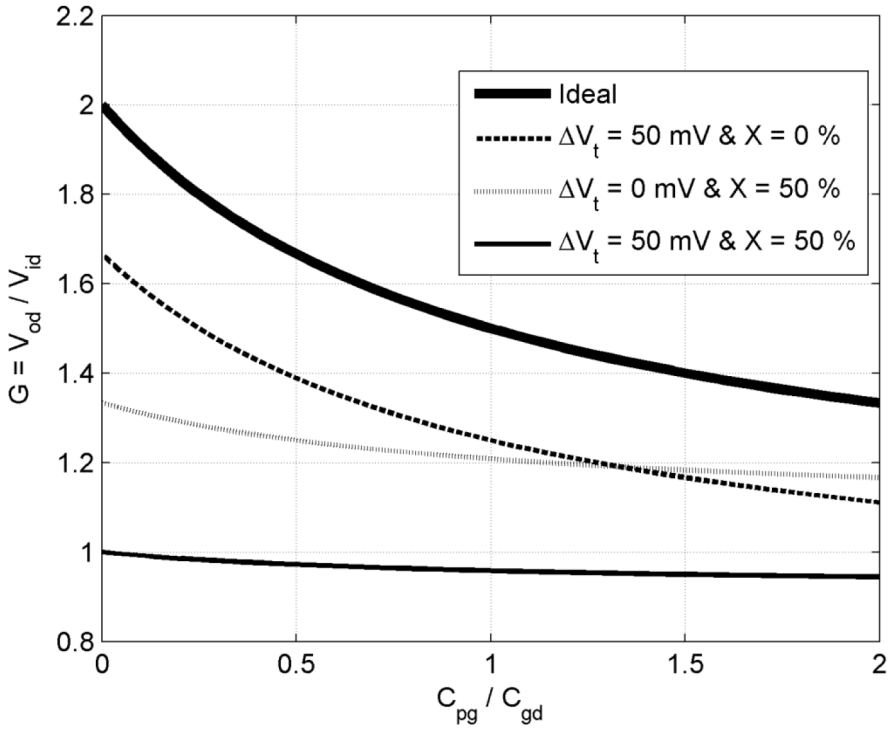


Figure 3.6 - Effects of parasitic capacitance, capacitor imbalance (*X*), and threshold (ΔV_t) imbalance on amplifier's gain

Threshold imbalance (the dashed curve) lowers the entire gain curve uniformly because it is amplified by the same G_{ideal} factor as the input signal. At $C_{pg} = C_{gd}$ in particular, a 50 mV threshold imbalance further reduces the parasitic limited G from 1.5 to less than 1.25. Capacitor imbalance also has similar reduction effect on gain, although less uniformly. Here, a 50 % capacitor imbalance, which can be caused by a large applied input voltage range, causes G at $C_{pg} = C_{gd}$ to drop from 1.5 to 1.2 (the gray curve). The combined effect of both threshold and capacitor imbalances (the thin dark solid curve) causes G to drop from 1.5 to 0.95, making this amplifier impractical. Therefore, general design guidelines to optimize the amplifier's performance should place a strong emphasis on minimizing the parasitic gate capacitance C_{pg}, the systematic threshold imbalance ΔV_t, and any variation of the moscaps with respect to the input signal (i.e. $X^s_{gsd} - X^a_{gd}$).

The most direct way to remedy gain limitations imposed by these bias-dependent imbalances is the use of an external capacitor C_{gs_ext}, as shown in Figure 3.7.

Adding this capacitor effectively enlarges C_{gs} and leads to a revised G_{ideal} given by (3.15). Sections 3.4 and Chapter 4 will explain the implications of adding this capacitance with more details in the context of linearity and settling performance.

$$G_{ideal} = \frac{C_{gsd}^{s} + C_{gs\_ext} + C_{pg}}{C_{pg} + C_{gd}^{a}} \tag{3.15}$$

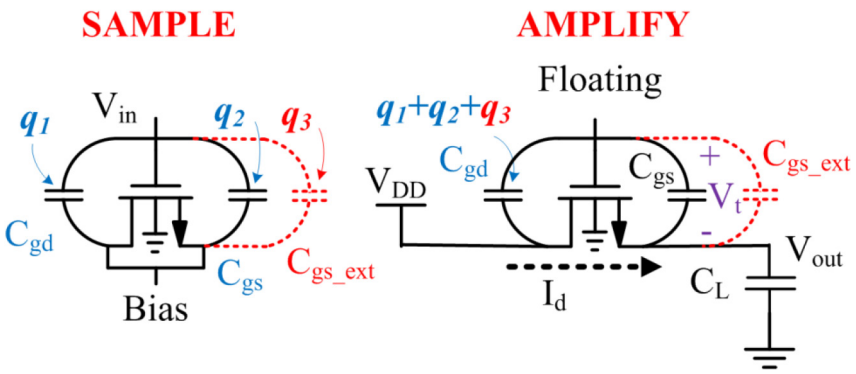


Figure 3.7 - Dynamic amplification with an added external cap C_{gs_ext} to boost gain

3.3. Amplifier's Output Common Mode

The previous section illustrates that both signal dependent capacitance and threshold imbalances cause uneven amount of charge to redistribute from the C_{gs} tubs to the C_{gd} tubs (see Figure 3.4 and Figure 3.5), altering the amplifier's differential output. These non idealities also affect the amplifier's common mode signal, further worsening its performance. In particular, using (3.8)-(3.13), Appendix A shows that the output common mode can be expressed as

$$V_{oc} = \frac{C_{pg} + \frac{C_{gsd}^{s} C_{px}}{C_{gsd}^{s} + C_{px}}}{C_{pg} + C_{gd}^{a}} V_{ic} + \frac{G_{ideal}}{4}\left(X_{gsd}^{s} - X_{gd}^{a}\right)\left(V_{id} - k\Delta V_{t}\right) - G_{ideal} k V_{t\_avg} + \frac{C_{gd}^{a}}{C_{pg} + C_{gd}^{a}} V_{DD} \tag{3.16}$$

Equation (3.16) shows that the output common mode is a complex function of the input common mode V_{ic}, the bias dependent capacitance imbalance $X_{gsd}^{s} - X_{gd}^{a}$, the average threshold voltage V_{t_avg}, and the supply voltage V_{DD}. Here, we are interested in

the amplifier's ability to reject undesired variations that can significantly perturb the output common mode. To this end, the following subsections discuss each of these four terms in detail.

3.3.1. Rejection of the Input Common Mode Variation

Our first concern is to examine the amplifier's sensitivity to variation in the input common mode. Taking the derivative of V_{oc} with respect to V_{ic} yields

$$G_{cm} = \frac{dV_{oc}}{dV_{ic}} = \frac{C_{pg} + \frac{C_{gsd}^{s} C_{px}}{C_{gsd}^{s} + C_{px}}}{C_{pg} + C_{gd}^{a}} \approx \frac{C_{pg} + C_{px}}{C_{pg} + C_{gd}^{a}} \tag{3.17}$$

One can also intuitively get to this expression by referring to Figure 3.2, which shows the equivalent CM half circuit of the amplifier during sampling. An input CM step ΔV_{ic} applied to the transistor's gate causes a voltage drop across C_{gsd}^{s} that is attenuated by the capacitive ratio given by (3.18). This results in a total charge stored on the gate given by (3.19).

$$\Delta V_{gs} = \frac{C_{px}}{C_{px} + C_{gsd}^{s}} \Delta V_{ic} \tag{3.18}$$

$$\Delta Q_g = C_{gsd}^{s} \Delta V_{gs} + C_{pg} \Delta V_{ic} \tag{3.19}$$

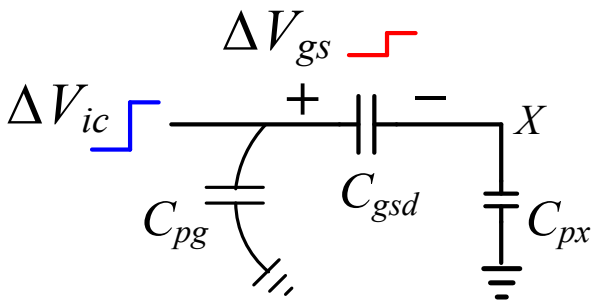


Figure 3.8 - Common mode equivalent circuit during the sampling phase

Transitioning from the sampling phase to the amplification phase, this total incremental CM charge stored on the gate redistributes itself across the sum of C_{gd} and C_{pg} to give an incremental output common mode voltage given by

$$\Delta V_{oc} = \frac{\Delta Q_g}{C_{pg} + C_{gd}^a} = \frac{C_{pg} + \frac{C_{gsd}^s C_{px}}{C_{gsd}^s + C_{px}}}{C_{pg} + C_{gd}^a} \Delta V_{ic} \tag{3.20}$$

As expected, the final expression of equation (3.20) is identical to that of (3.17). For the amplifier to reject variation in the input common mode, the common mode gain G_{cm} must therefore be minimized. As seen Figure 3.8, this condition is equivalent to minimizing both the parasitic capacitance of the gates (C_{pg}) and of node X (C_{px}). For the amplifier to have any rejection at all, however, the common mode gain G_{cm} must be less than unity. If we assume that $C_{gsd}^s \gg C_{px}$ as in (3.17), this condition leads to a requirement that $C_{px} < C_{gd}^a$; i.e. the parasitic capacitance at the bottom plate of the sampling cap (node X) must be less than the amplifier's gate-to-drain capacitance.

From Figure 3.2, it can be seen that C_{px} is the total parasitic capacitance of the drain and source nodes. Considering the three possible implementations using Bulk CMOS as shown in Figure 3.9, this parasitic capacitance is the junction capacitance between (a) the drain/source diffusions and the p-substrate, (b) the drain/source diffusions and the p-body, or (c) the p-body and the n-well. Due to their large areas of contact, these parasitic capacitances can easily exceed the fringing gate-to-drain capacitance C_{gd}. Simulation of a 90-nm Bulk CMOS implementation configured as shown in Figure 3.9c returns a $G_{cm} > 1.3$. Thus, for a nominal common mode voltage of 0.6 V (i.e. half of V_{DD}), a 100 mV input common mode fluctuation would rail the ADC after 6 stages. For the implementations in Figure 3.9a and b, the common mode gain is G_{cm} is slightly lower but still generally larger than unity. As explained in more detail in Section 3.4, fluctuation in the output common mode in general is undesirable, since it perturbs the bias point of the transistor, reducing its available output range as a result. These reasons therefore discourage the implementation of this particular amplifier topology using a Bulk CMOS process.

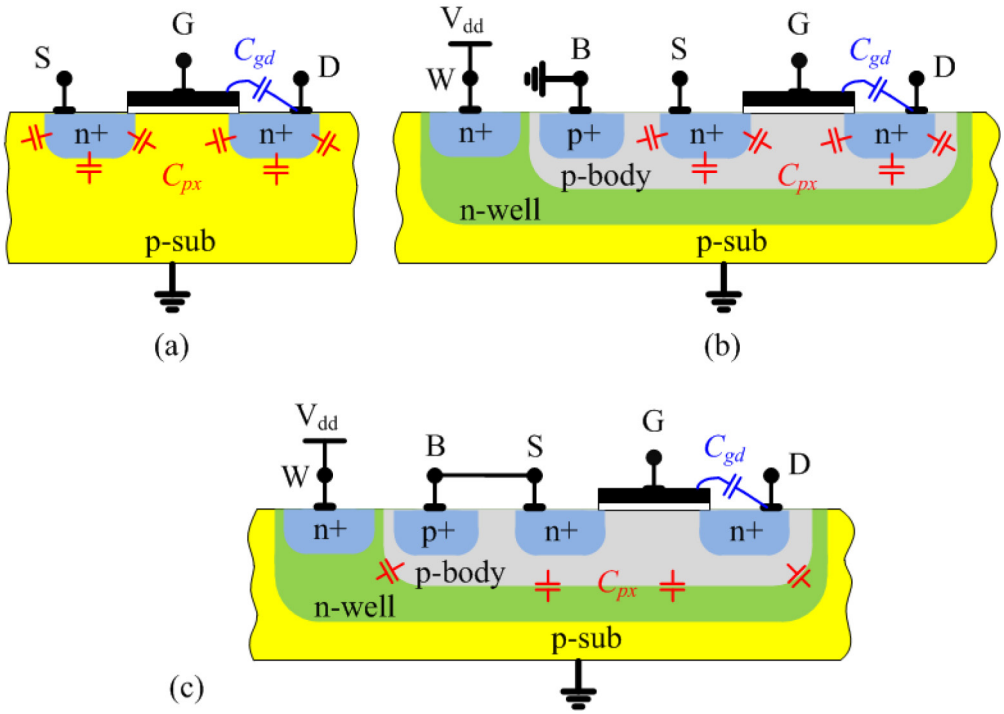


Figure 3.9 - Bulk CMOS implementation with (a) a conventional process, (b) a twin-well process with a grounded body, and (c) a twin-well process with the body tied to source

For a thin-film SOI process with a grounded body as shown in Figure 3.10, C_{px} is the sidewall junction capacitance between the drain/source diffusions and the body. This capacitance is typically much less than the gate-to-drain capacitance C_{gd} for two reasons. The first is due to the much smaller area of contact (i.e. the junction sidewalls) between the drain/source and the body, and the second is due to the reversed bias normally applied on these junctions during amplifier's operation. As an illustration, simulation of a 65-nm grounded-body SOI CMOS implementation with a junction reversed bias of ~ 0.45 V and $C_{pg} = C_{gd}^{a}$ shows a $G_{cm} < 0.6$ (which is close to the minimum $G_{cm\_min} = 0.5$ obtained from (3.17) when $C_{px} = 0$).

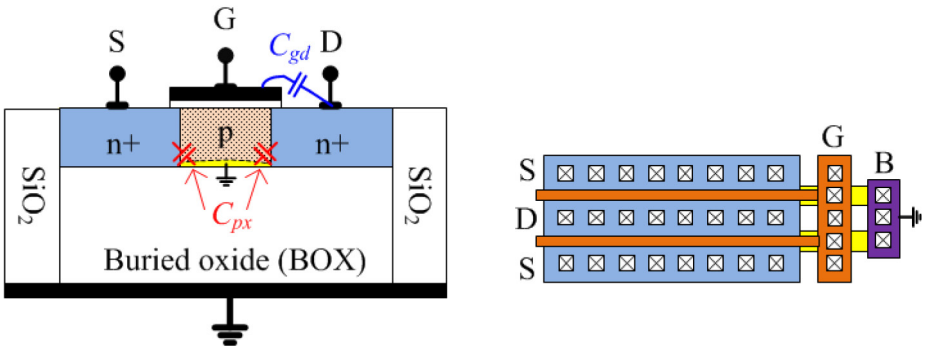


Figure 3.10 - SOI CMOS implementation with a grounded body: cross section and layout views

Without a grounded body, C_{px} becomes just the series combination of the junction capacitance between the drain/source diffusions and the body and the oxide capacitance from the body to ground, which is very small for a typical SOI BOX thickness and can thus be neglected in (3.17). The common mode gain G_{cm} of a floating body SOI MOSCAP is therefore smaller (i.e. closer to 0.5) and only limited by the gate parasitic capacitance.

3.3.2. Differential to Common Mode (DM-CM) Conversion

Signal dependent capacitance imbalances, which show up as the second term $X^s_{gsd} - X^a_{gd}$ in (3.16), cause DM-CM conversion that also perturbs the output common mode. Due to the charge redistribution nature of the amplifier, such DM-CM conversion is also amplified by the amplifier's gain, enhancing the perturbation on the output common mode. This is similar to the effects that the variation in the widths of the tubs has on the differential signal, as discussed in Section 3.2. As an illustration, for a gain of $G_{ideal} = 2$, an applied input range of V_{id_peak} = 600 mV, and a threshold imbalance of ΔV_t = 50 mV, a capacitive imbalance of $X^s_{gsd} - X^a_{gd}$ = 50% would lead to a ~140 mV variation in the output common mode, reducing the available output range of the amplifier by the same amount. Minimizing capacitive imbalances is therefore essential in mitigating such (signal dependent) CM perturbation.

3.3.3. REJECTION OF SUPPLY VARIATION

Finally, the output common mode V_{oc} is also sensitive to variation in the average threshold voltage V_{t_avg} and the supply V_{DD}, as depicted by the third and fourth terms in (3.16). However, since DIBL and finite output resistance affect the two circuit halves in opposite fashions, an increase or decrease in V_{tp} is cancelled to some extent by a corresponding decrease or increase in V_{tm}, keeping their average value V_{t_avg} relatively constant. Simulation shows that its variation is small and can generally be neglected. On the other hand, the output common mode V_{oc} is a strongly sensitive function of supply variation. The common mode power supply rejection ratio (PSRR) is defined as

$$PSRR_{CM} = \frac{dV_{DD}}{dV_{oc}} = \frac{C_{pg} + C_{gd}^{a}}{C_{gd}^{a}} \tag{3.21}$$

For an example case of $C_{pg} = C_{gd}^{a}$, any supply variation, though attenuated by ~50% through this capacitive ratio, couples directly into the amplifier's output common mode. A large supply variation therefore would significantly perturb the amplifier's common mode output, reducing the available output swing and worsening the amplifier's signal-to-noise ratio (SNR). Here, a higher $PSRR_{CM}$ requires a larger gate parasitic capacitance C_{pg}, but this requirement conflicts with the condition for high signal gain and low common mode gain [see (3.14) and (3.17)]. To maintain a good SNR, therefore, a clean supply and/or calibration of the output CM (with respect to V_{DD} variation) are necessary.

3.4. SAMPLING LINEARITY

Section 3.2 discusses the impact of signal dependent imbalances on the amplifier's gain. This section focuses on the amplifier's linearity. In general, two primary sources of distortion exist to degrade the amplifier's performance. The first is the nonlinearity of the moscap as a sampling capacitor during the sampling phase, and the second stems from the transistor's nonlinear characteristics as a current source during the amplification phase. This section focuses on sampling linearity, leaving the

latter topic to Chapter 4 which will be discussed in the context of the amplifier's settling performance.

From Figure 3.2 which shows the equivalent sampling network, the differential sampling capacitor with an added linear external cap C_{gs_ext} for extra gain is given by (3.22).

$$C_S = C_{gs} + C_{gd} + C_{pg} + C_{gs\_ext} \tag{3.22}$$

$$\frac{\Delta C_S}{C_S} = \frac{\Delta(C_{gs} + C_{gd} + C_{pg})}{C_{gs} + C_{gd} + C_{pg} + C_{gs\_ext}} \tag{3.23}$$

For an implementation with a partially depleted (PD) SOI CMOS process, C_S strongly depends on how the body is biased. If the body is left floating as shown in Figure 3.11a, its potential fluctuates greatly with the gate voltage due to coupling via the gate oxide capacitance. The width of the depletion region inside the body would vary as a result, causing the associated body-to-drain and body-to-source junction capacitances (i.e. C_{bd} and C_{bs}) to also fluctuate. Unless the transistor is biased in strong accumulation or strong inversion (i.e. where a layer of charge underneath the gate oxide serves as a shield), such variation in junction capacitances distorts the effective C_{gs} and C_{gd} seen by the input driver, leading to a nonlinear and signal dependent C_S.

Biasing the transistor in either accumulation or strong inversion, however, cause implementation difficulties that are discussed in more detail in Chapter 4 and 5. Alternatively, (3.23) suggests that we can linearize C_S by arbitrarily enlarging a linear C_{gs_ext} but, as will be apparent in Chapter 4, such drastic increase would lead to a dramatic cost in speed. Even with a small signal swing and operating at a bias point with the widest possible linear range, the combination of degraded performance (i.e. lower speed and resolution) resulting from such capacitance variation makes this amplifier unattractive and encourages the finding of other design alternatives.

Not sure what amplifier you are referring to — the one w/ external C_S, or all dynamic amps ...

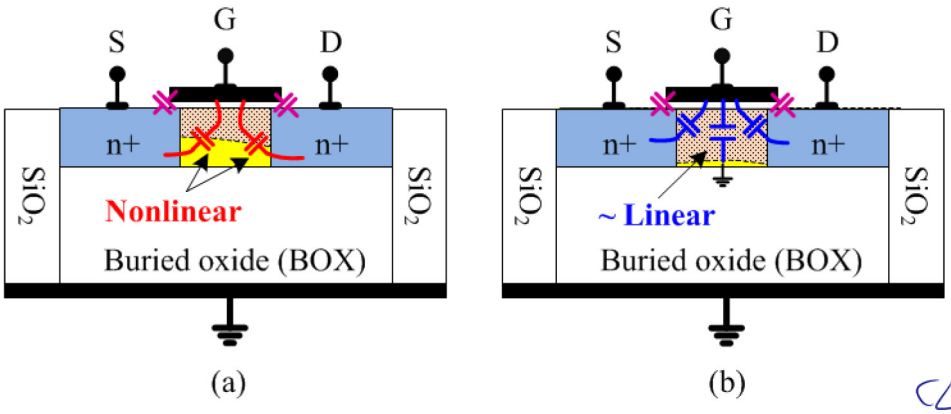


Figure 3.11 - SOI MOSCAP with (a) a floating body and (b) a grounded body

Fortunately, a solution to our problem comes from a slight change in the SOI moscap configuration. To be more specific, grounding the body of a PD SOI transistor, as shown in Figure 3.11b, essentially removes most of the variation and nonlinearity associated with C_S. As the body is grounded, this region becomes almost entirely depleted, and its width stays relatively constant and independent of the applied gate voltage. This leads to two important benefits. The first is a much more linear C_S seen by the input driver, significantly improving the amplifier's sampling linearity. The second is the resulting decoupling of the amplifier's bias point from linearity considerations. In other words, since C_S remains relatively constant regardless of the transistor's bias point, we now have freedom in choosing any bias point that optimizes other performance aspects of the amplifier (i.e. speed, common mode, etc.). The applied signal range is, to first order, no longer restricted by the linear operating range of the MOSCAP, leading to an enhanced signal-to-noise-and-distortion ratio (SNDR) of the amplifier.

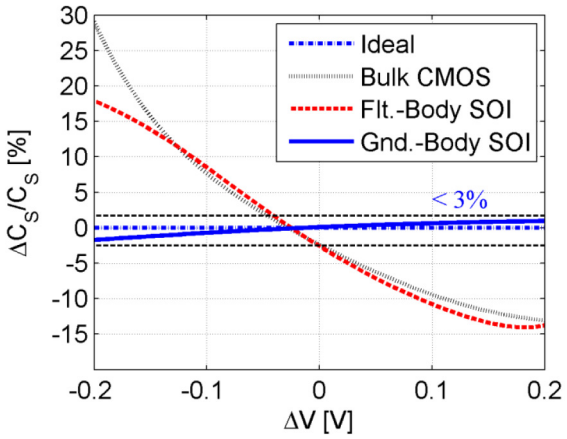
To illustrate the signal dependence of the sampling capacitance C_S, we plot the normalized variation of C_S (i.e. $\Delta C_S/C_S$) with respect to an applied input voltage ΔV in Figure 3.12. Here, it is important to distinguish two important effects associated with this C-V variation. The first is the large signal variation that results in a linear gain reduction as discussed in Section 3.2. To separate this linear-reduction effect from linearity consideration, a straight-line fit is constructed for each C-V curve of Figure

3.12. The deviations of these curves from their respective straight-line fits are then computed and plotted in Figure 3.13 to provide a measure of their nonlinearity.

As can be seen in Figure 3.12 and Figure 3.13, a 65-nm PD SOI MOSCAP with a floating body, represented by the red dashed line, shows a large signal C-V variation of greater than 30 % and a nonlinear deviation of ~5 %, leading to a corresponding simulated sampling linearity of 4…6 bits. On the other hand, grounding the body of the same MOSCAP (the blue solid curve) drastically reduces both the large signal variation and nonlinear deviation to ~3 % and 0.5 %, respectively. This corresponds to a simulated sampling linearity of 8...9 bits, a dramatic improvement over the former case. The C-V curve of the 90-nm Bulk CMOS implementation is also included here for comparison to show both a greater large signal variation of ~45 % and nonlinear deviation of ~10 %. From the perspective of both gain (see Section 3.2) and sampling linearity, a thin film SOI process with a grounded body is therefore much better suited for dynamic amplification than its Bulk counterpart.

The improvement that comes with grounding the SOI body, however, does not come without penalty. In particular, this implementation requires an extension of the body diffusion beyond the gate which in turn needs to be contacted, as depicted by the layout view in Figure 3.10. This introduces additional parasitic gate-to-body capacitance that enlarges C_{pg}, reducing the amplifier's gain and worsening its common mode rejection according to (3.15) and (3.17). To minimize this parasitic contribution, the body contacts are thus pulled slightly away from the gate. However, the increased IR drop that comes with a longer diffusion region causes the body to be less of a perfect ground, which in turn can degrade the linearity of C_S and potentially introduce memory effects at high conversion rates. Nevertheless, simulation shows that such high order effects only result in a minor impact on the amplifier's performance, and the 3...4-bit improvement in linearity from grounding the transistor's body far exceeds these drawbacks.

I am confused, since the resistance of this body in the wide device you show is very large. Do you know its RC time?



at what bias point?

Figure 3.12 - Variation in sampling capacitance vs. the applied input voltage

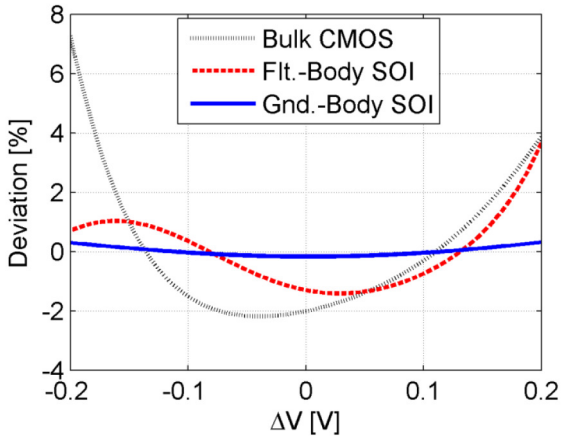


Figure 3.13 - Deviation of capacitance variation $\Delta C_S/C_S$ from a straight line fit

3.5. NOISE ANALYSIS

The noise contribution of each pipelined stage after the SHA comes from two main sources. The first is the noise from the sampling phase, and the second is from the DAC and CM capacitors that interface the summing node. Since these 2 noise sources are uncorrelated, we can analyze them independently and sum the results later. Denoting α^i as the scaling factor of stage i, the noise voltage acquired during sampling phase by stage i is approximated by (see Appendix C)

$$v_{ns}^2(i) \approx \frac{k_B T}{\alpha^i C_S} \tag{3.24}$$

During amplification, this voltage noise is amplified by the amplifier's gain G and appears at the output as

$$v_{na1}^2(i) = \frac{k_B T}{\alpha^i C_S} G^2 \tag{3.25}$$

Independently of this noise source, it can be shown that the noise charge contribution on the summing node from resetting the DAC and CM capacitors (see Chapter 5) can be approximated by

$$q_{nx}^2(i) = k_B T \alpha^i (C_{dac} + C_{cm}) \left(\frac{C_S}{C_{dac} + C_{cm} + C_S} \right)^2 \tag{3.26}$$

Referring this noise charge to the output, we have

$$\begin{aligned} v_{na2}^2(i) &= \frac{q_{nx}^2(i)}{\alpha^{2i} (C_{pg} + C_{gd})^2} = \frac{k_B T}{\alpha^i} \frac{C_{dac} + C_{cm}}{(C_{pg} + C_{gd})^2} \left(\frac{C_S}{C_{dac} + C_{cm} + C_S} \right)^2 \\ &= \frac{k_B T}{\alpha^i (C_{dac} + C_{cm})} G_{ideal}^2 \left(\frac{C_{dac} + C_{cm}}{C_{dac} + C_{cm} + C_S} \right)^2 \end{aligned} \tag{3.27}$$

The total voltage noise at each stage's output at the end of amplification is therefore the sum of the above two sources and given by

$$v_{nout}^2(i) = v_{na1}^2 + v_{na2}^2 = \frac{k_B T}{\alpha^i} \left[\frac{G^2}{C_S} + \frac{G_{ideal}^2}{C_{dac} + C_{cm}} \left(\frac{C_{dac} + C_{cm}}{C_{dac} + C_{cm} + C_S} \right)^2 \right] \tag{3.28}$$

In a differential implementation, the noise contribution of stage i is twice the noise of each individual amplifier. Referring this total output noise to the ADC input, the noise contribution stage i is therefore

$$v_n^2(i) = 2 \frac{v_{nout}^2(i)}{G^{2i}} \tag{3.29}$$

Assuming that the ADC has a front-end SHA with a gain of G that precedes the n pipelined stages, the total input referred noise of the ADC is given by the sampling noise of the SHA and the input referred noise of all the stages that follow

$$v_{n\_tot}^2 = 2\frac{k_B T}{C_S} + \sum_{i=1}^{n} v_n^2(i) \tag{3.30}$$

$$v_{n\_tot}^2 = 2\frac{k_B T}{C_S} + 2k_B T\left[\frac{G^2}{C_S} + \frac{G_{ideal}^2}{C_{dac}+C_{cm}}\left(\frac{C_{dac}+C_{cm}}{C_{dac}+C_{cm}+C_S}\right)^2\right]\sum_{i=1}^{n}\left(\frac{1}{\alpha G^2}\right)^i$$

$$v_{n\_tot}^2 = 2\frac{k_B T}{C_S}\left[1 + (1+M)\frac{1}{\alpha G^2}\frac{1-\left(\frac{1}{\alpha G^2}\right)^n}{1-\frac{1}{\alpha G^2}}\right] \tag{3.31}$$

What is this?

$$where\ M = \frac{C_S}{C_{dac}+C_{cm}} G_{ideal}^2 \left(\frac{C_{dac}+C_{cm}}{C_{dac}+C_{cm}+C_S}\right)^2$$

For a stage scaling factor of 1, (3.31) reduces to

$$v_{n\_tot}^2 = 2\frac{k_B T}{C_S}\left\{1 + \frac{1+M}{G^2}\cdot\frac{1-G^{-2n}}{1-G^{-2}}\right\} \tag{3.32}$$

As will be discussed in Section 4.3, G is approximated to be ~1.55 to 1.65. Using all other relevant values from Chapter 5 (i.e. C_S ~ 55 fF, C_{cm} ~ 10 fF, C_{dac} ~ 7 fF, etc.), (3.32) yields an estimated total ADC input referred noise of 800…850 µV-rms.

3.6. Chapter Summary

This chapter analyzed the practical design tradeoffs of important parameters of the amplifier such as gain, linearity, common mode and supply rejection, and noise. The discussion combined the use of a hydraulic model and analytical results to offer both an intuitive and quantitative view of these tradeoffs. In particular, the discussion emphasized the importance of minimizing the signal dependence of the moscap and reducing key parasitic capacitances to achieve the highest linearity and gain while ensuring adequate common mode rejection. It was also concluded from the discussion that a thin-film SOI process with a grounded body offered the best performance combination of gain, linearity, and common mode rejection. In the next two chapters, the settling performance and implementation details of the amplifier will be discussed, demonstrating how the discussed advantages of the SOI process also help increase the amplifier's speed.

CHAPTER 4. INCOMPLETE SETTLING

The previous chapters presented the concept and tradeoffs of dynamic amplification while neglecting the transient or settling behavior of the amplifier. This chapter discusses this topic in details. Specifically in Section 4.1, we study the process of charge transfer from C_{gs} to C_{gd} of a single-ended amplifier by examining the settling trend of its gate-to-source voltage (V_{gs}). This analysis is carried out using basic models of the transistor in both strong and weak inversion to highlight the key parameters that dictate the V_{gs} settling. Following this is a study of the settling trend of the stage's differential output residue V_{od}, which will be shown in Section 4.2 to depend on both the V_{gs} settling and the transistor's transconductance g_m. From this discussion, incomplete settling will be introduced in Section 4.3 as a method to drastically increase the amplifier's settling speed. Along with a comparator look-ahead scheme that will be discussed in Chapter 5, this technique, which is implemented in a 65-nm SOI process, provides a 10x improvement in conversion rate over prior art [11]. Finally, Section 4.5 summarizes of the chapter.

4.1. THE DYNAMIC CHARGE TRANSFER PROCESS AND V_{GS} SETTLING

A general limitation of this amplifier is its slow transient response due to a decreasing drain current as the transistor approaches its final and steady state. Referring to Figure 4.1 which shows the transient progression of the charge transfer from C_{gs} to C_{gd} of a single-ended amplifier during amplification, it can be seen that the speed of this redistribution process is the rate of decaying of V_{gs}. To understand this V_{gs} settling behavior, we consider the transistor in amplification as shown in Figure 4.2, which is the differential version of the circuit in Figure 3.3.

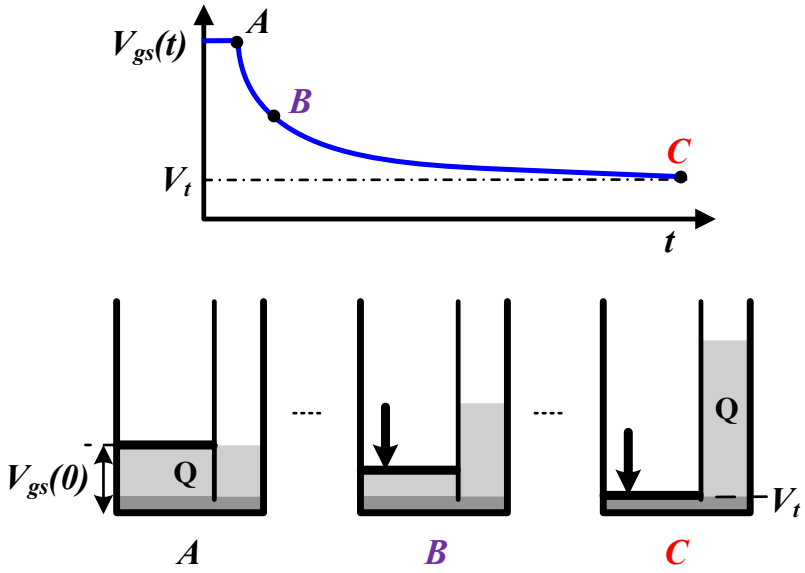


Figure 4.1 – The transient progression of the charge transfer process from C_{gs_tot} to C_{gd}

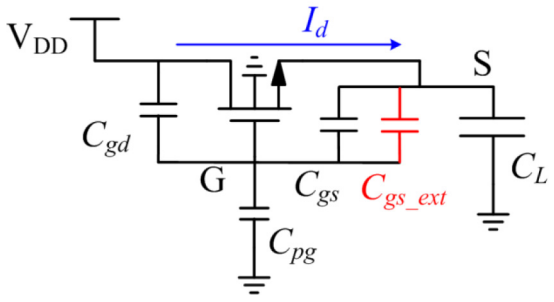


Figure 4.2 – Equivalent circuit of the dynamic amplifier during amplification

Here, we are interested in the maximum charge transfer speed of the amplifier, which is set by the rate of discharging $C_{gs\_tot} = C_{gs} + C_{gs\_ext}$ (or dV_{gs}/dt). Thus,

$$f_{qxfer} \propto -\frac{dV_{gs}}{dt} = \frac{dV_s}{dt} - \frac{dV_g}{dt} \tag{4.1}$$

Since the gate is coupled with the source via C_{gs_tot}, we thus have

$$\frac{dV_g}{dt} = \frac{C_{gs} + C_{gs\_ext}}{C_{gs} + C_{gs\_ext} + C_{pg} + C_{gd}} \frac{dV_s}{dt} \tag{4.2}$$

Therefore

$$f_{qxfer} \propto -\frac{dV_{gs}}{dt} = \frac{C_{pg} + C_{gd}}{C_{gs} + C_{gs\_ext} + C_{pg} + C_{gd}} \frac{dV_s}{dt} = \frac{1}{G_{ideal}} \frac{dV_s}{dt} \tag{4.3}$$

$$= \frac{1}{G_{ideal}} \frac{I_d}{C_{Ltot}}$$

$$f_{qxfer} G_{ideal} \propto \frac{I_d}{C_{Ltot}} \approx \frac{I_d}{\alpha C_S} \tag{4.4}$$

The above equations are general, because no assumptions have been made regarding the transistor's characteristics. Since the charge transfer speed is set by the decaying rate of V_{gs} (dV_{gs}/dt), which is proportional to the load charging rate (dV_s/dt), it is intuitively clear that the larger the drain current I_d, the faster the transfer speed and vice versa. This is confirmed by (4.4) which shows that the product of the charge transfer speed f_{qxfer} and voltage gain G_{ideal} is proportional to the drain current I_d and inversely proportional to the load size (~ αC_S). For a fixed drain current and load capacitance, this product is constant; indicating that a faster charge transfer speed requires a lower gain.

As charge is transferred from C_{gs} to C_{gd} during amplification, however, the drain current I_d decays rapidly over time due to a proportional reduction of V_{gs} as demonstrated in Figure 4.1. This current decaying process slows down the load charging rate (dV_s/dt), reducing dV_{gs}/dt, and explains for the reason why it takes considerably more time for V_{gs} to progress from *B* to *C* than from *A* to *B*. Consequently, if we could afford to discard all the charge transferred between B and C and still have sufficient voltage amplification, the amplifier's timing overhead would

be small since it is only the (short) time interval between A and B. This is the basis of our incomplete settling technique that will be discussed in more detail in Section 4.3.

In the context of complete settling (complete charge transfer), however, for a noise limited amplifier ($SNR \propto C_S$ - see Section 3.5), increasing the resolution by one bit quadruples C_S. For an amplifier with fixed current and gain, equation (4.3) predicts that each additional bit causes a 4x reduction in speed. Continuing our analysis by assuming a pipelined stage scaling factor of α, the load capacitance can be approximated as $C_{Ltot} \approx \alpha C_S = \alpha\left(C_{gs} + C_{gs\_ext} + C_{pg} + C_{gd}\right)$. Equation (4.3) thus becomes

$$f_{qxfer} \propto -\frac{dV_{gs}}{dt} \propto \frac{1}{G_{ideal}} \frac{I_d}{\alpha C_S} = \frac{C_{pg} + C_{gd}}{\alpha {C_S}^2} I_d \tag{4.5}$$

If the design objective is to maximize the amplifier's gain by increasing C_{gs_ext}, equation (4.5) predicts that this would result in a quadratic reduction in transfer speed (i.e. each doubling of gain causes a 4x reduction in speed). If C_S is distortion limited, as discussed in Section 3.4, adding C_{gs_ext} has the effect of linearizing C_S according to (3.23). For instance, each doubling of C_S, via adding an appropriate C_{gs_ext}, reduces the relative distortion (i.e. $\Delta C_S/C_S$) by a factor of two leading to an increase of 1 bit in sampling linearity. The cost, however, is again a 4x reduction in transfer speed according to (4.5).

In summary, the above qualitative discussion has shown that a high charge transfer speed requires a large drain current and a small load capacitance. It was also concluded that improving either the amplifier's resolution or voltage gain tends to cause a significant reduction in charge transferring speed. In the following discussion, quantitative results are developed to further understand these tradeoffs. Specifically, we analyze the V_{gs} transient behavior using both the strong and weak inversion models of the transistor and compare the results against simulation. The purpose of these analyses, however, is not to describe the transient behavior exactly, but to gain key insight into the amplifier's settling trends confirming our previous intuition and to

motivate the use of incomplete settling (in Section 4.2) as a means to drastically improve the amplifier's speed.

4.1.1. V_{GS} SETTLING IN STRONG INVERSION

Before analyzing the settling behavior of the amplifier, a current source I_b is added to Figure 4.2 to give Figure 4.3. As will be clear in the following discussion, the purpose of this current source is to bias the transistor in a way that improves its settling characteristics. Using the square-law model, neglecting finite output resistance, and by referring to Figure 4.3, we have

$$\frac{dV_s}{dt} = \frac{I_d - I_b}{C_{Ltot}} = \frac{K(V_{gs} - V_{th})^2 - I_b}{C_{Ltot}} = \frac{KV_{ov}^2 - I_b}{C_{Ltot}} \tag{4.6}$$

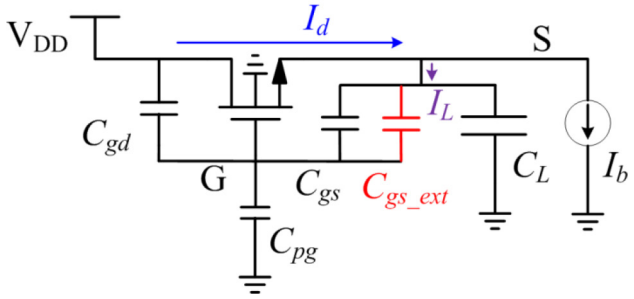


Figure 4.3 – Dynamic amplifier during amplification with an added current source I_b

Here, the effective threshold voltage V_t is a function of the output potential V_s (due to DIBL and finite output resistance as previously discussed in Section 3.2) and hence varies with time. However, simulation shows that we can approximate V_t variation as a proportional function of the output's variation. In other words, we can write

$$\frac{dV_t}{dt} \approx k_D \frac{dV_s}{dt} \quad where\ k_D = const. \tag{4.7}$$

For the 90-nm Bulk CMOS process used in this study as an example, $k_D \sim 4\%$. Towards the end of amplification t_s and as V_s settles to its final value, we expect that the drain current approaches the value of the bleed current I_b. Assuming that the

transistor does not go into weak inversion (i.e. $V_{gs} \geq V_t$), we can write $I_d(t_s) = I_b = K[V_{ov}(t_s)]^2$, leading to

$$\frac{dV_{gs}}{dt} = -\frac{1}{G_{ideal}}\frac{dV_s}{dt} = -\frac{1}{G_{ideal}}\frac{K[V_{ov}^2(t) - V_{ov}^2(t_s)]}{C_{Ltot}} \tag{4.8}$$

From the derivation in Appendix C, the solution to the above differential equation is

$$V_{gs}(t) \approx V_{gs}(0) - V_{ov}(t_s) \cdot \left(\frac{2}{1 - \frac{R_f + 1}{R_f - 1} e^{t/\tau}} + R_f - 1 \right) \tag{4.9}$$

$$where \quad R_f = \frac{V_{ov}(0)}{V_{ov}(t_s)} \quad and \quad \tau \approx G_{ideal}\frac{C_{Ltot}}{2KV_{ov}(t_s)} = G_{ideal}\frac{C_{Ltot}}{g_m(t_s)}$$

In (4.9), k_D and K are constants that depend on process parameters. The settling behavior thus depends solely on the ratio of overdrive voltages R_f, the load capacitance ($C_{Ltot} \approx \alpha C_S$), and the transistor's transconductance at time t_s. In a typical pipelined ADC implementation, where the voltage gain and load capacitance αC_S are fixed by design, the settling time evolution of $V_{gs}(t)$ is thus primarily dictated by τ, which is inversely proportional to the transistor's final transconductance $g_m(t_s)$. Since this transconductance is set by I_b, increasing I_b, therefore, speeds up the V_{gs} decaying rate.

As an illustration, Figure 4.4 plots (4.9) along with the simulated transient response of a design example. The parameters used to construct the model curve in this figure are taken from simulated data ($K \approx 4.2\ mA/V^2$, $k_D \approx 4\%$, $C_{Ltot} \approx 55\ fF$, $G_{ideal} \approx 2.86$, $V_{ov}(0) \approx 120\ mV$, and $V_{ov}(t_s) \approx 70\ mV$). The bleed current is set high (I_b ~ 18.5 µA) to keep the transistor in strong inversion.

what's the device?

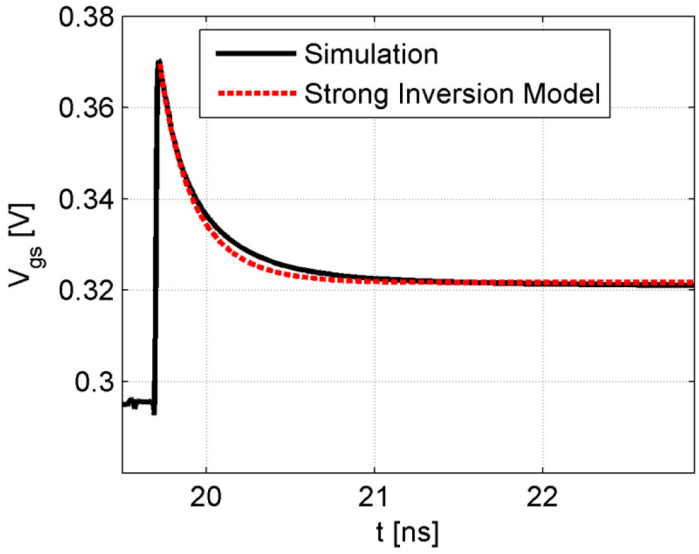


Figure 4.4 - V_{gs} settling trend in strong inversion

Figure 4.4 shows that our square-law model's results agree reasonably well with simulation, although the former predicts a higher decaying rate of V_{gs}. In general, small discrepancies exist between the results obtained from this basic model and simulation. Nevertheless, these differences do not affect the main conclusions of this analysis. To proceed further, let us define the normalized V_{gs} error $\varepsilon(t)$, which is the difference between the value of V_{gs} at time t and its final value $V_{gs}(\infty)$, as follows

$$\varepsilon(t) = \frac{V_{gs}(t) - V_{gs}(\infty)}{V_{gs}(0) - V_{gs}(\infty)} = \frac{2}{(R_f + 1)e^{t/\tau} - (R_f - 1)} \tag{4.10}$$

Using this definition, a complete charge transfer process at time t_s requires that $\varepsilon(t_s) \approx 0$. Looking at this from another perspective, for a given V_{gs} error spec ε_{spec}, the minimum settling time t_s can be derived from the above equation and plotted to give

$$t_s = \tau \cdot ln\left(\frac{R_f - 1}{R_f + 1} + \frac{2}{\varepsilon_{spec}}\frac{1}{R_f + 1}\right) \tag{4.11}$$

Assuming a constant τ (i.e. constant I_b), Figure 4.5 plots the V_{gs} error versus settling time for different R_f values (i.e. different initial overdrive $V_{ov}(0)$ values). It is apparent that a smaller ε is equivalent to a larger t_s (i.e. it takes longer to transfer more charge from C_{gs} to C_{gd}). However, it is important to note that, for a given ε target, the

final settling result stays within a time constant τ, regardless of the value of R_f. In other words, even with a 10x variation in the initial overdrive or R_f, the difference in settling time t_s for $R_f = 2$ and $R_f = 20$ is less than 0.8τ according to Figure 4.5. The settling time therefore shows a weak dependence on the applied input and depends primarily on the final transconductance value (that is set by the bleed current I_b) with an average slope of ~ 1 decade per τ (dec/τ).

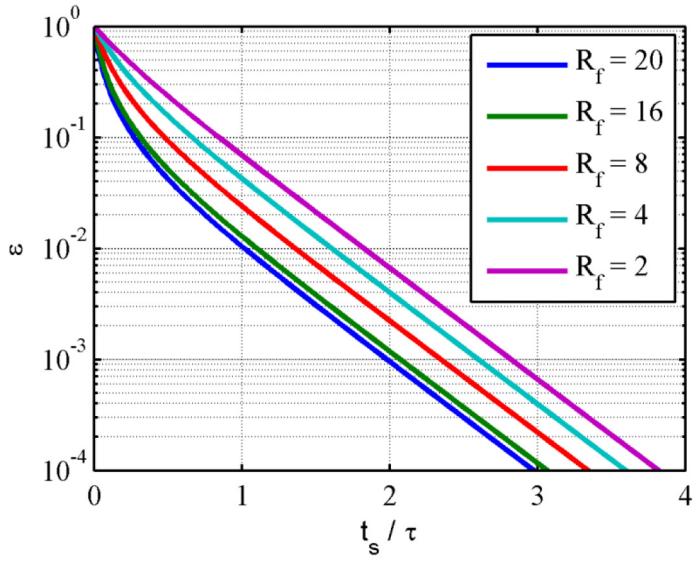


Figure 4.5 - V_{gs} settling error vs. normalized settling time in strong inversion

Looking at Figure 4.5 from a different perspective reveals another crucial observation concerning the amplifier's linearity. At a fixed settling time t_s, a different input signal value (i.e. different R_f) corresponds to a different accuracy level. For instance, a large applied input, corresponding to $R_f = 20$, settles to a 0.1 % error at $t_s/\tau = 2$. On the other hand, a small applied input, corresponding to $R_f = 2$, settles to only ~1 % at the same t_s/τ. This therefore suggests a nonlinear relationship between the V_{gs} settling error and the applied input, as can be subsequently confirmed in Figure 4.6, which plots ε vs. R_f for different values of t_s/τ. From observing Figure 4.6, one may speculate that terminating the settling process too soon can cause significant distortion to the amplifier's output settling (V_s), because it is a proportional function of V_{gs} [see (4.8)]. More settling time is therefore needed if such nonlinearity is to be mitigated. Section 4.3.2 will look at this issue in more detail.

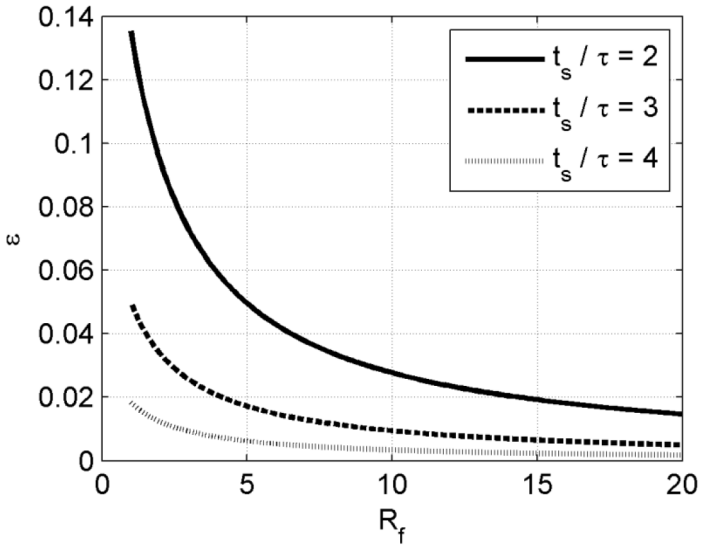


Figure 4.6 – V_{gs} settling error vs. different ratio of overdrive voltages in strong inversion

4.1.2. V_{GS} SETTLING IN WEAK INVERSION

At some point during the settling process, if I_b is sufficiently small, the transistor enters weak inversion, where V_{gs} becomes less than the extrapolated threshold voltage V_t. The transistor thus operates more closely to an npn bipolar transistor (BJT) [28] leading to

$$\frac{dV_s}{dt} = \frac{I_d - I_b}{C_L} = \frac{I_o exp\left(\frac{V_{gs}}{nV_T}\right) - I_b}{\alpha C_S} \tag{4.12}$$

$$where \quad V_T = \frac{k_B T}{q}$$

Solving the above differential equation in Appendix D to get

$$V_{gs}(t) = -nV_T ln\left(\frac{I_o}{I_b} - R_o \exp\left(-{}^{t}/_{\tau}\right)\right) \tag{4.13}$$

$$where \quad R_o = \frac{I_o}{I_b} - \exp\left(-\frac{V_{gs}(0)}{nV_T}\right) \quad and \quad \tau = G_{ideal}\frac{C_{Ltot}}{g_m(t_s)}$$

Repeating the same exercise as before, we plot (4.13) with (4.9) on the same graph as our simulated results. This time, I_b is set small for both equations. Figure 4.7

shows that the square law equation |fails to the track simulated curve as the transistor approaches and enters weak inversion. Using data estimated from simulation ($C_{Ltot} \approx 55\ fF$, $G_{ideal} \approx 2.86$, $I_b \approx 1.8$ µA, and $n \approx 1.5$), it can be seen that the weak inversion curve tracks the simulated response reasonably well, demonstrating that (4.13) can also serve as a guide to improve the charge transfer speed in weak inversion.

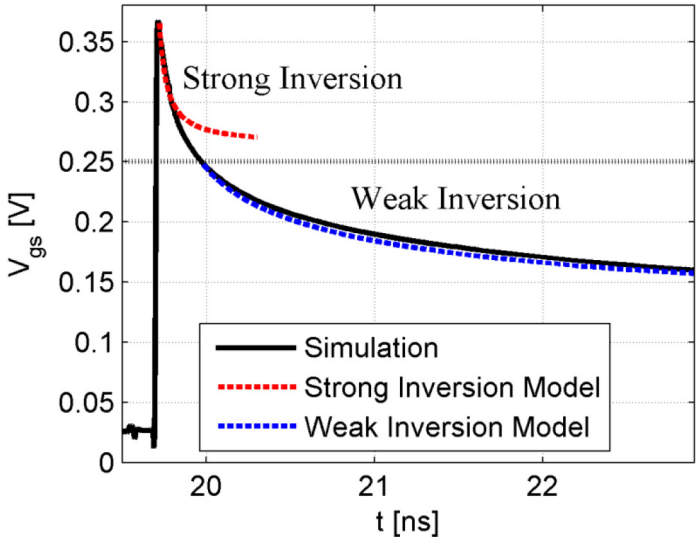


Figure 4.7 – V_{gs} settling trend in weak inversion

Following the same development steps as before, the normalized V_{gs} error $\varepsilon(t)$ is thus

$$\varepsilon(t) = \frac{V_{gs}(t) - V_{gs}(\infty)}{V_{gs}(0) - V_{gs}(\infty)} = -\frac{ln\left(1 - R_1 \exp\left(-t/\tau\right)\right)}{\frac{V_{gs}(0)}{nV_T} + ln\left(\frac{I_o}{I_b}\right)} \tag{4.14}$$

$$where \quad R_1 = 1 - \frac{I_b}{I_o}\exp\left(-\frac{V_{gs0}}{nV_T}\right) \quad and \quad \tau = G_{ideal}\frac{C_{Ltot}}{g_m(t_s)}$$

Much larger than before, right?

Plotting the above V_{gs} error versus normalized settling time, similar to Figure 4.5, for various initial values V_{gs0} leads to Figure 4.8. From this figure, one again draws similar conclusions as before. In particular, the charge transfer time is inversely proportional to the V_{gs} error with a reduced average slope of less than 0.5 dec/τ. This indicates a weaker drive capability of the transistor in weak inversion. Moreover, for a fixed gain G_{ideal} and load αC_S, the charge transfer time is primarily dictated by the

magnitude of I_b and is more dependent on the transistor's initial conditions than when in strong inversion (i.e. at any given settling error ε, the settling time t_s varies by about one time constant τ for various V_{gs0} values).

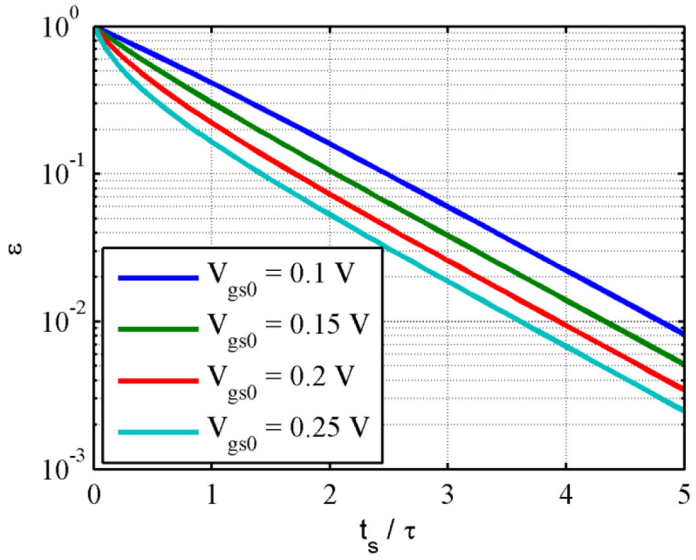


Figure 4.8 - V_{gs} settling error vs. vs. normalized settling time in weak inversion

In terms of linearity, Figure 4.9 shows similar characteristics with respect to the applied input. However, by allowing more time for the transistor to settle, such nonlinearity behavior can be mitigated significantly. Therefore, this presents a tradeoff between linearity and speed of the amplifier: less time allocated for settling corresponds to a higher distortion.

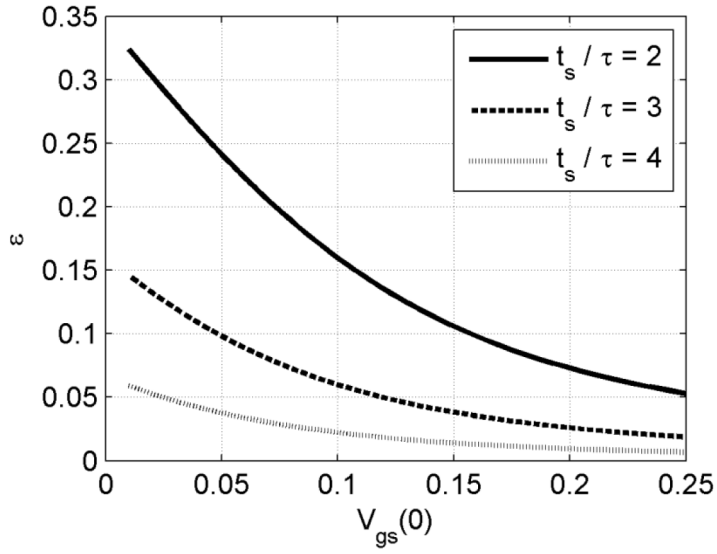


Figure 4.9 - V_{gs} settling error vs. initial V_{gs} value in weak inversion

4.1.3. SUMMARY OF V_{GS} SETTLING

The above analyses demonstrated that, for a fixed voltage gain and load capacitance, the time it takes to transfer charge from C_{gs} to C_{gd} of the amplifier is primarily dictated by the transistor's final transconductance value, which is set by the current I_b. The highest possible transfer speed therefore requires a large I_b. Next, it was also shown that there exists a nonlinear relationship between the V_{gs} settling error and the magnitude of the applied input signal. This V_{gs} nonlinearity can only be mitigated by allocating more time for the amplifier to settle, presenting a tradeoff between linearity and speed for the amplifier. In the following section, we examine the transient behavior of the amplifier's differential output residue V_{od} and show how these observations can help maximize its settling speed.

4.2. SETTLING OF THE DIFFERENTIAL OUTPUT VOLTAGE

To study the settling behavior of the differential amplifier's differential output V_{od} and leverage the observations made in Section 4.1, Figure 4.10 shows the transient simulation results of the circuit in Figure 3.3, using I_b as a sweeping parameter. As can be seen from this figure, when I_b is increased, not only does the amplifier's V_{gs} but also

its V_{od} settle more quickly. Without I_b, the amplifier never fully settles, and its V_{gs} continues to decay slowly and indefinitely (or its output continues to rise indefinitely until the transistor's output reaches V_{DD}). Though settling more quickly, the amplifier with a higher I_b has a smaller V_{oc} rising slope (due to an increasing I_b that reduces I_L – see Figure 4.3) and a larger final V_{gs} value (indicating a smaller amount of charge transferred from C_{gs} to C_{gd}).

This settling behavior of the differential output residue V_{od} can be understood in two steps. The first is to visualize that the effect of I_b is to modify the effective V_t's of the transistors, with a larger V_t corresponding to a higher I_b and vice versa. Ignoring bias dependent imbalances, (3.4) demonstrates that the gain is independent of V_t, and thus the settled differential output remains the same regardless of the value of I_b. This is visible in Figure 4.10 which shows that all the V_{od} curves converge to the same final value as t goes to infinity.

Next, to understand why I_b helps speed up the amplifier's settling, we refer to Figure 3.3 and consider the following expressions

$$\left|\frac{dV_{od}}{dt}\right| = \left|\frac{dV_{op}}{dt} - \frac{dV_{om}}{dt}\right| = \frac{\Delta I_L}{C_L} \tag{4.15}$$

$$\frac{dV_{oc}}{dt} = \frac{1}{2}\left(\frac{dV_{op}}{dt} + \frac{dV_{om}}{dt}\right) = \frac{I_{L\_avg}}{C_L} \tag{4.16}$$

The above equations express the speed of the amplifier's differential and common mode outputs, dV_{od}/dt and dV_{oc}/dt, as functions of the output settling rates of the individual half circuits, dV_{op}/dt and dV_{om}/dt. From considerations of V_{gs} settling in the previous section, it was concluded that the highest charge transfer speed between C_{gs} and C_{gd} requires the largest current that can be delivered to the load during amplification. For the differential amplifier of Figure 3.3, this requirement is equivalent to maximizing the load currents of both circuit halves, which is also the same as maximizing their average load current I_{L_avg}. This condition, however, only speeds up the amplifier's output common mode V_{oc} and does not necessarily lead to a high settling speed of the differential mode V_{od}, which is proportional only to the difference in load currents ΔI_L of the two circuit halves.

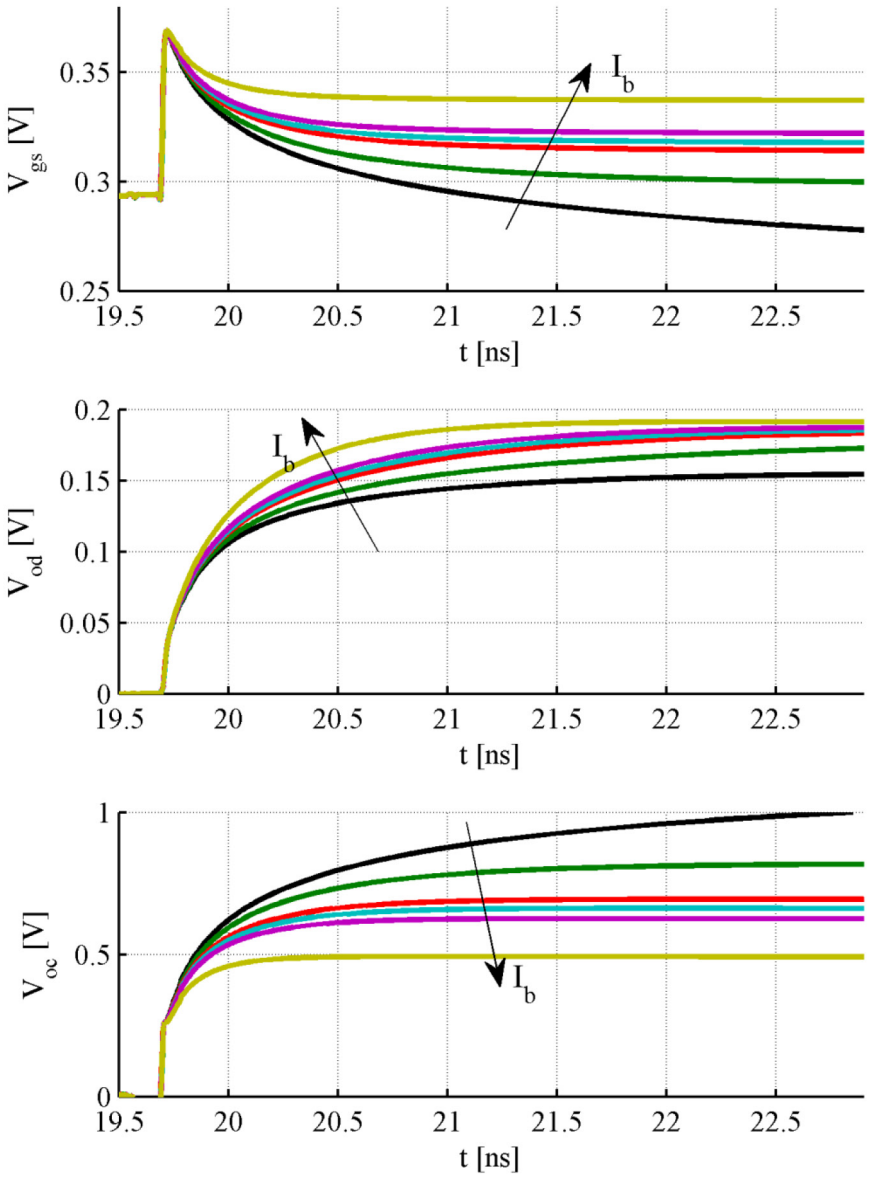


Figure 4.10 - Simulated settling behavior of the amplifiers. Here, only V_{gs} of the positive half is shown, and the arrow indicates the direction of increasing I_b (≈ 0 to 40 µA).

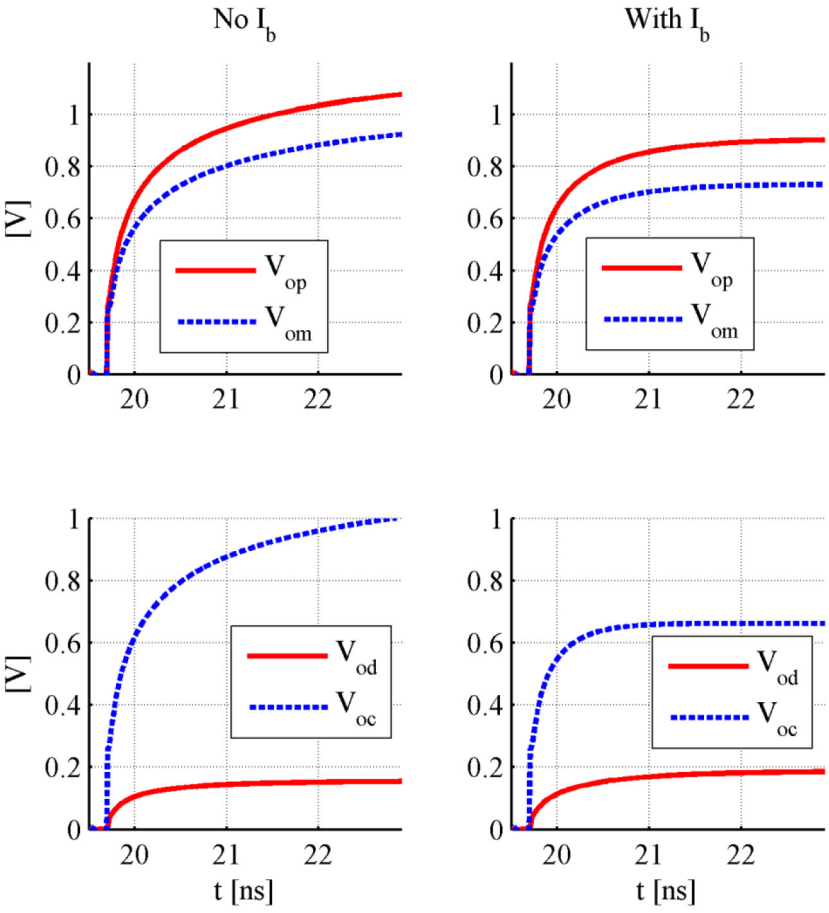


Figure 4.11 - Common mode and differential mode settling

During amplification, since both outputs V_{op} and V_{om} rise monotonically in the same direction to their final values, our intuition tells us that the differential output rate dV_{od}/dt, due to being the difference, is typically slower than the common mode rate dV_{oc}/dt. This can be seen by examining Figure 4.11 which plots the transient settling of the differential circuit of Figure 4.3 with and without I_b, showing in both cases that the common mode rate dV_{oc}/dt is greater than the differential rate dV_{od}/dt. The differential mode rate, however, exceeds the common mode rate only if one circuit half is significantly faster than the other. [In fact, (4.15) and (4.16) show that dV_{od}/dt increases and exceeds dV_{oc}/dt if one circuit half is at least three times faster than the other.] Due to the signal dependence and decaying nature of the amplifier's current during amplification (see Section 4.1), this condition typically cannot be met (or is met only for a very brief moment with a sufficiently large input). The outputs of the two circuit

halves therefore rise at similar rates, and the differential mode rate dV_{od}/dt is always slower than the common mode rate dV_{oc}/dt.

Consequently, if the highest settling speed is to be achieved, we must maximize not the average load current delivered to the two circuit halves but their differential load current ΔI_L during amplification. The bleed current accomplishes this objective at the cost of increased power consumption. To understand why, it is helpful to examine the I-V characteristic of the transistor shown in Figure 4.12. Due to the nonlinear nature of the transistor, varying the bleed current essentially slides the transistor bias point along the I-V curve. As a result, a higher bias point (i.e. point B) would have a higher I-V slope (i.e. higher transconductance g_m) than the lower bias point (i.e. point A). For a fixed differential input voltage V_{id}, this higher g_m at B thus yields a higher ΔI_d, leading to a higher differential load current ΔI_L and resulting in an increased dV_{od}/dt and a faster settling time as apparent in Figure 4.10 when I_b is increased. Expressing this mathematically and denoting the transconductance of the positive and negative half circuits as g_{m_pos} and g_{m_neg}, we have

$$\Delta I_L \propto g_{m\_pos}V_{ip} - g_{m\_neg}V_{im} = g_{m\_avg}V_{id} + \Delta g_m V_{ic} \tag{4.17}$$

Ignoring the signal dependent term Δg_m in (4.18), it is clear that an amplifier with the highest (average) transconductance g_{m_avg}, (set by I_b) offers the highest speed. As I_b or g_{m_avg} is increased, however, static power dissipation also increases, and the amplifier approaches class-A behavior. Furthermore, as can be seen in Figure 4.10, increasing I_b also causes the output common mode V_{oc} to decrease due to an increase in the required V_{gs} to maintain the high static current. This may impact the output signal swing since it is a function of the common mode output. These power saving and common mode considerations therefore place an upper bound on the maximum bleed current that can be used in a real design.

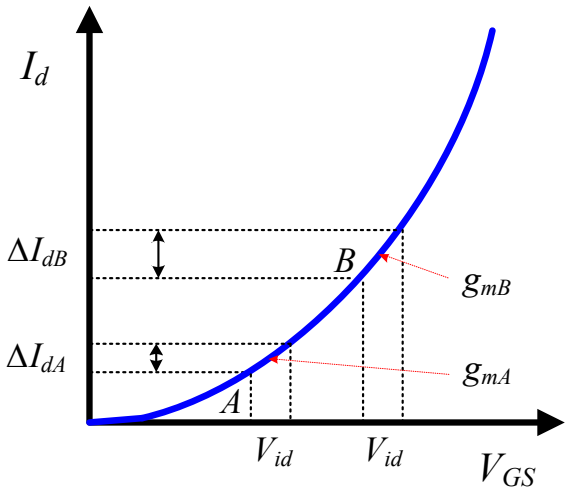


Figure 4.12 - Effect of I_b on ΔI_d

4.3. INCOMPLETE SETTLING

The previous discussion shows that that the use of the bleed current I_b speeds up the settling speed of the amplifier's differential output V_{od}. However, such improvement in speed is offset by an overhead increase in static power consumption. To simultaneously maintain low power and high speed, one must be able to keep the transistor in strong inversion (i.e. high g_m) without relying on a large bleed current. This is accomplished via the use of a large external capacitance C_{gs_ext} (see Figure 4.3).

Typical output settling curves of an amplifier with and without I_b look similar to the solid and dotted V_{od} curves shown in Figure 4.13, where it takes a certain settling time t_{s2} or more for the output to completely reach its desired value V_{od_f}. The presence of a bleed current places a lower bound on V_{gs} forcing the transistor to stay in the high g_m region of the I-V curve burning static current to improve its settling speed over the case without I_b. In contrast to this scheme, the second method discussed here replaces I_b with a large C_{gs_ext} across the gate and source terminal of the transistor (see Figure 4.2). This effectively enlarges the circuit time constant and thus reduces the decaying rate of V_{gs}, keeping the transistor in the high g_m region longer to give the same effect of maximizing the differential load current ΔI_L as a high I_b.

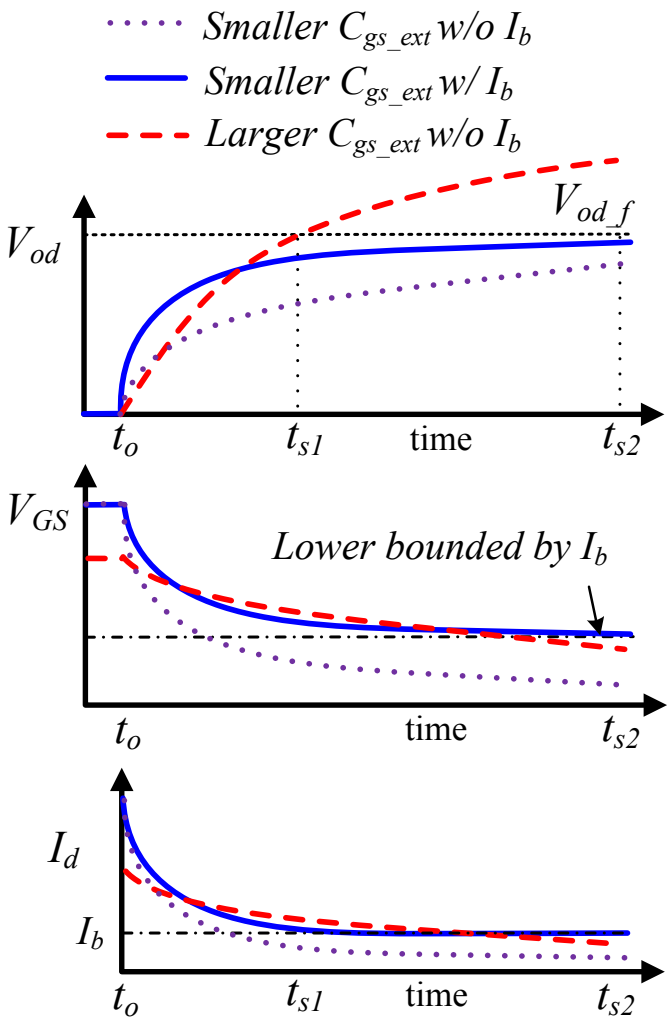


Figure 4.13 – Incomplete settling

The difference with the latter scheme, however, is that the amplifier's gain G_{ideal} now also increases as a result due to an enlarged C_{gs_tot} (see Figure 3.7), forcing the transistor to settle to a higher output value. Nevertheless, since the amplifier only needs to reach V_{od_f}, the amplification phase can be terminated early at time t_{s1}, leading to a significant reduction in settling time (i.e. $\Delta t_s = t_{s2} - t_{s1}$). Furthermore, this method of using a large C_{gs_ext} also costs significantly less total power than the I_b scheme as can be seen in the I_d plot of Figure 4.13. In particular, the integrated drain current of the solid curve from t_o to t_{s2} is larger than that of the dashed line from t_o to t_{s1}. Also looking at this from the efficiency perspective, all of the current in the large C_{gs_ext} scheme goes

directly to the load, while the I_b scheme is less efficient because much of its current is wasted.

Although increasing C_{gs_ext} improves the settling speed, diminishing return and even longer settling time result once C_{gs_ext} exceeds a certain value. For a fixed transistor size and hence fixed current drive strength, increasing C_{gs_ext} increases in the amplifier's load capacitance C_{Ltot} and also causes the circuit time constant τ to become excessively large, reducing the speed of both the V_{gs} decaying and the output settling as a result (see (4.4), (4.9), and (4.13)). As an illustration, Figure 4.14 shows the simulated output response of an ideal amplifier with increasing values of C_{gs_ext}. In this figure, the gray curve is the amplifier's input with a dashed vertical marker to denote the sampling instant during the sampling phase. The other dark curves correspond to the amplifier's output response during amplification for various values of C_{gs_ext}. For a fixed settling budget as marked in the figure, an optimal C_{gs_ext} in the vicinity of ~1.5 to 2.2 C_{gg} yields the highest gain. Beyond these optimal values, the output begins to show slower ramp rates for increasing C_{gs_ext} corresponding to a lower amplifier gain for the same settling budget.

The optimal choice of C_{gs_ext} depends on a combination of design tradeoffs between noise, power, speed, and parasitic capacitance. Budgeting for thermal noise ($\propto kT/C_S$) places a lower bound on the total sampling capacitance C_S (equal to C_{gs} + C_{gd} + C_{gb} + C_{gs_ext}). For a given total C_S, the optimal ratio of C_{gs_ext} / C_{gg} for maximum gain at t_s is found through iterations and in simulation as shown in Figure 4.14. Next, parasitic capacitances (estimated at the beginning and more precisely determined after layout extraction) tend to reduce the maximum achievable gain and shift the optimal ratio to a higher value. In this work, by targeting a settling time t_s of ~700 ps for a 500 MS/s converter, it was found through design iterations and after layout extraction that the optimal region for the ratio C_{gs_ext} / C_{gg} is shallow with an optimum value of ~3 as seen in Figure 4.15. This choice leads to an estimated post-layout extraction gain of 1.55…1.65.

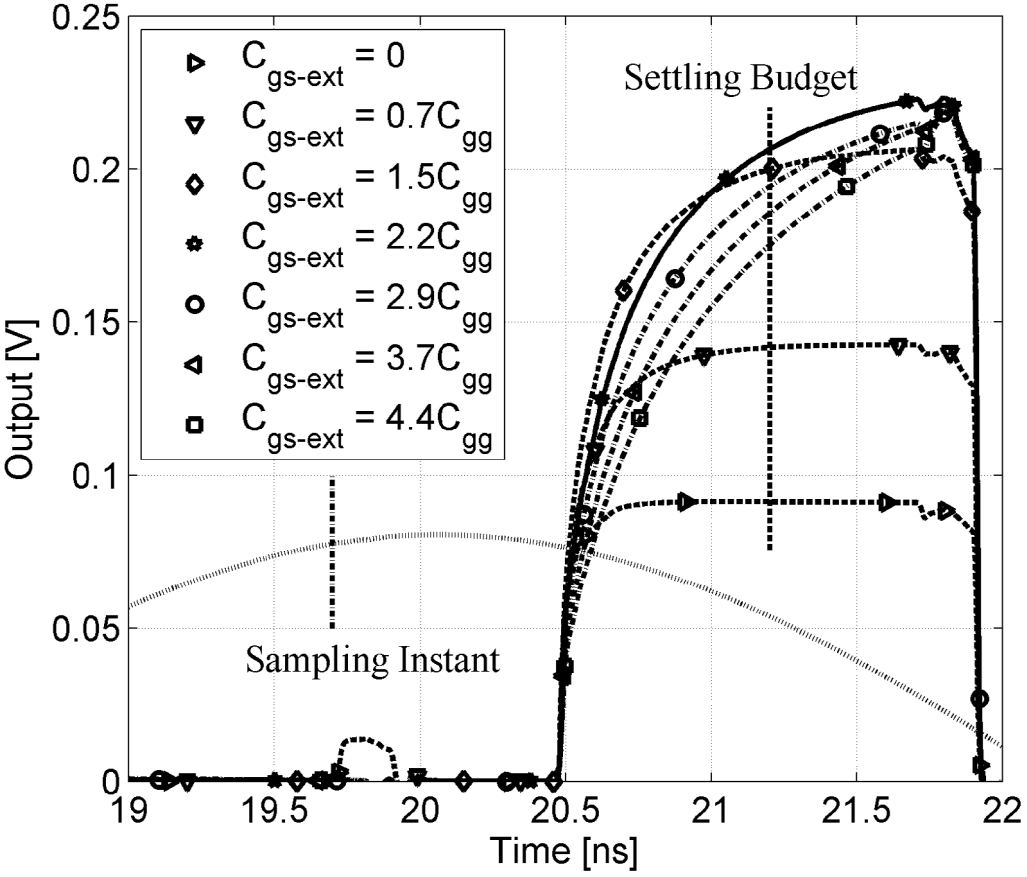


Figure 4.14 - Effect of increasing C_{gs_ext} on amplifier's output response.

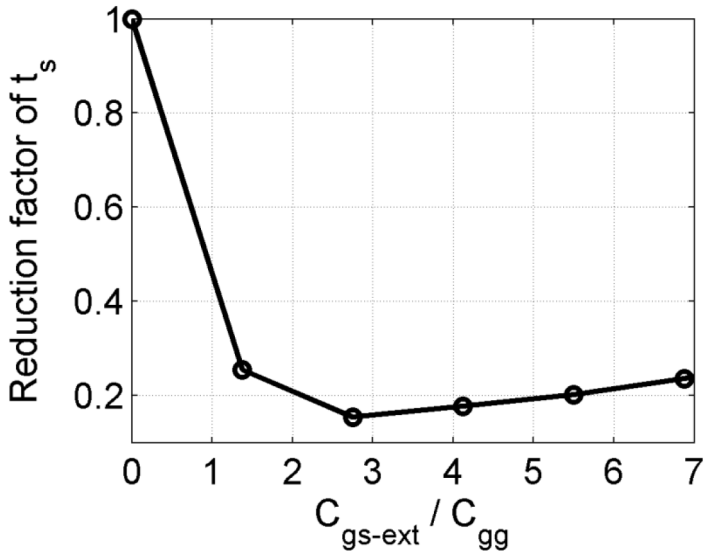


Figure 4.15 - Reduction factor of settling time via incomplete settling with C_{gs_ext}

4.3.1. JITTER SENSITIVITY DUE TO INCOMPLETE SETTLING

Although increasing the external capacitance C_{gs_ext} helps increasing the amplifier's speed, the incompletely settled output of the amplifier has a non-zero slope that becomes susceptible to clock jitter. The less settled the amplifier, the higher the output slope. Given a clock jitter variance of δ_t^2 , the jitter induced voltage noise of the amplifier with an output slope S is given by

$$v_{nj}^2 = S^2\delta_t^2 < \left(\frac{v_{out\_max}}{t_s}\right)^2 \delta_t^2 \tag{4.19}$$

In the above equation, v_{out_max} is the maximum value of the amplifier's output, and t_s is the settling budget allocated to the amplifier. Since the transient output of the amplifier is always convex, the term on the right of (4.19) serves as a conservative upper bound on the noise induced by clock jitter. Figure 4.16 plots this jitter noise v_{nj} against δ_t for various amplifier output slopes.

From Figure 4.16, it is clear that a higher degree of incomplete settling corresponds to more noise induced by clock jitter. Jitter noise therefore places an upper bound on how fast (or how incompletely settled) the amplifier can be. As an illustration, for an effective ADC resolution of seven bits or better, a 10 ps-rms clock jitter dictates that the amplifiers settle with an output slope of ~0.1 mV/ps to maintain a jitter noise level roughly the same as the quantization noise level (i.e. ~1 mV-rms). Along with gain considerations, this also places a lower bound on t_s according to (4.19) limiting the maximum achievable conversion rate as a result. In this work, simulation shows that by allocating an amplifier's settling time of t_s ~ 700 ps, the output slope is ~0.1 mV/ps, corresponding to an estimated jitter noise voltage of 0.1…1 mV-rms for a clock jitter of 1…10 ps-rms.

4.3.2. NONLINEARITY DUE TO INCOMPLETE SETTLING

As previously shown in Figure 4.6 and Figure 4.9, an early termination of the amplification phase leads to an incomplete charge transfer process that results in a nonlinear relationship between the V_{gs} settling errors ε and the applied input signal. To

see how this can impact the performance of the amplifier, the linearity of the SOI amplifier's differential output V_{od} is evaluated with respect to the amplifier's degree of incomplete settling. To be more specific, by applying a sinusoidal signal to the amplifier's input and then taking the Fast Fourier Transform (FFT) of its sampled output, the amplifier's linearity can be evaluated by subsequently observing its spurious free dynamic range (SFDR). In particular, the amplifier's output SFDR is plotted against the settling time offset to give Figure 4.17.

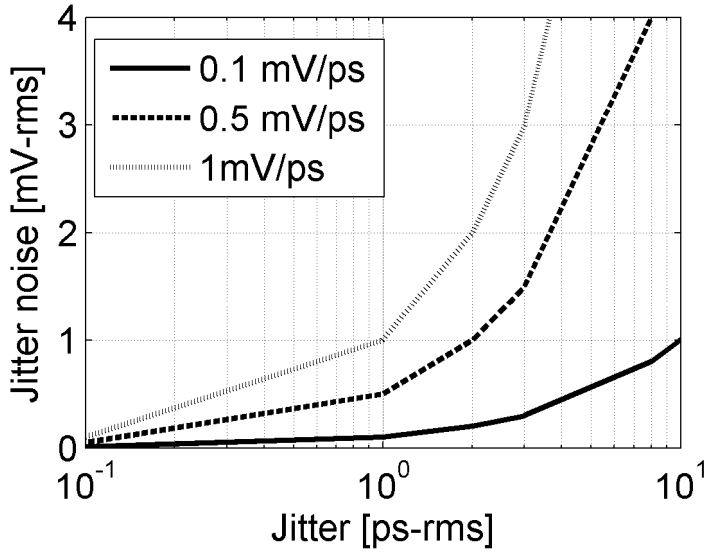


Figure 4.16 – Jitter induced noise for various output slopes

In Figure 4.17, a zero timing offset denotes near complete settling of the amplifier (i.e. $t_s \geq 3.2$ ns), while a negative timing offset signifies incomplete settling (i.e. $t_s < 3.2$ ns). As the timing offset becomes more negative, the amplifier's SFDR degrades significantly, placing a lower limit on the minimum allowable t_s. To keep distortion sufficiently low for an 8-bit converter (i.e. SFDR > 50 dB), the timing offset must be higher than roughly –2.5 ns, corresponding to a minimum allocated t_s on the order of ~700 ps.

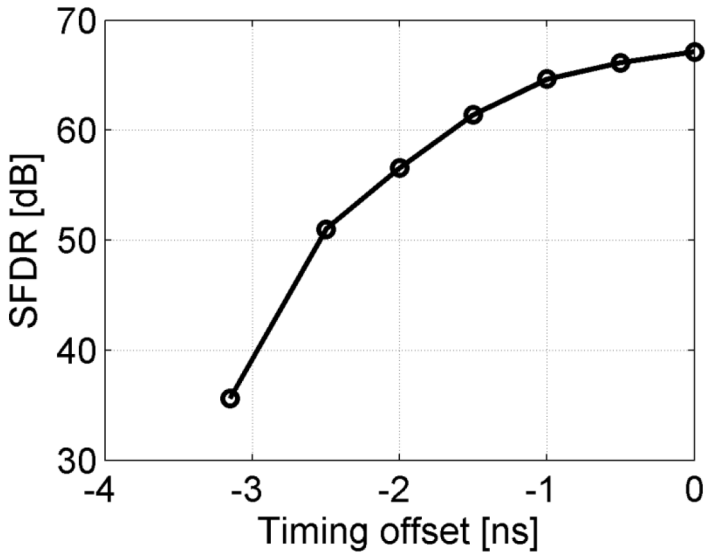


Figure 4.17 – Effect of incomplete settling on amplifier's linearity

4.4. CHAPTER SUMMARY

This chapter discussed the settling performance of the dynamic amplifier during amplification. It was shown that maximizing the settling speed is equivalent to maintaining a high amplifier's transconductance. Two techniques were thus proposed to accomplish this. The first was the use of a large external static current to bias the amplifier in strong inversion thereby keeping its transconductance high. The other was through an increase in the amplifier's gain via an external gate capacitance. Combined with incomplete transient settling, the latter method achieves an even higher settling speed and without the need for a static current. Incomplete settling, however, introduces additional jitter-induced noise and distortion that limit the maximum achievable conversion rate. In the next chapter, we will show how this technique can be combined with a comparator look-ahead scheme to build an ADC with a conversion rate of 500 MS/s.

CHAPTER 5. ARCHITECTURAL DESIGN

This statement doesn't really say anything & usually can be deleted

In this chapter, we will describe how the understanding of the design tradeoffs developed in Chapter 3 and 4 led to an amplifier implementation suitable for a high speed pipelined ADC. The chapter is organized as follows. We begin with a description of the ADC's overall architecture in Section 5.1, followed by a discussion of the clock design and the chip's data acquisition flow in Section 5.2. Next, the implementation of the ADC's sample-and-hold (SHA) and stages are presented in Section 5.3 and 5.4, respectively. In Section 5.4, we will describe a comparator look-ahead scheme that was used along with the incomplete settling technique of Chapter 4 to achieve a conversion rate of 500 MS/s. The layout strategy is then discussed in Section 5.5, followed by the chapter summary in Section 5.6.

5.1. PROTOTYPE ADC ARCHITECTURE

From the analysis of Chapter 4 which shows that the amplifier's speed is inversely proportional to its gain, the 1-bit-per-stage architecture becomes the most

suitable choice for a high-speed ADC due to its low gain requirements. Targeting a conversion rate of 500 MS/s, the conversion cycle T_s is limited to 2 ns. Due to the synchronous operation of the pipelined ADC (see the review in Chapter 2), the sample phase Φ_2 and amplify phase Φ_4 are each allocated ~30…35 % of the cycle time (i.e. ~650…700 ps). Since a bias or reset phase Φ_1 and a CM adjust phase Φ_3 (equal Φ_1) are also needed for a complete cycle of operation (more in Section 5.3), the ADC clocking scheme thus consists of 4 primary non-overlapping phases: bias, sample, adjust, and amplify.

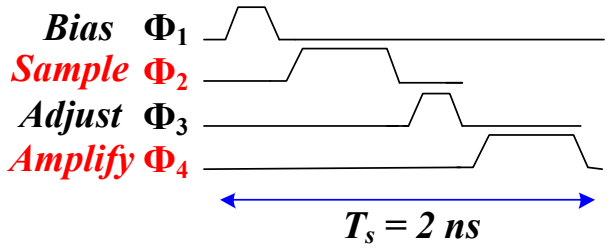


Figure 5.1 - Amplifier's clocking scheme

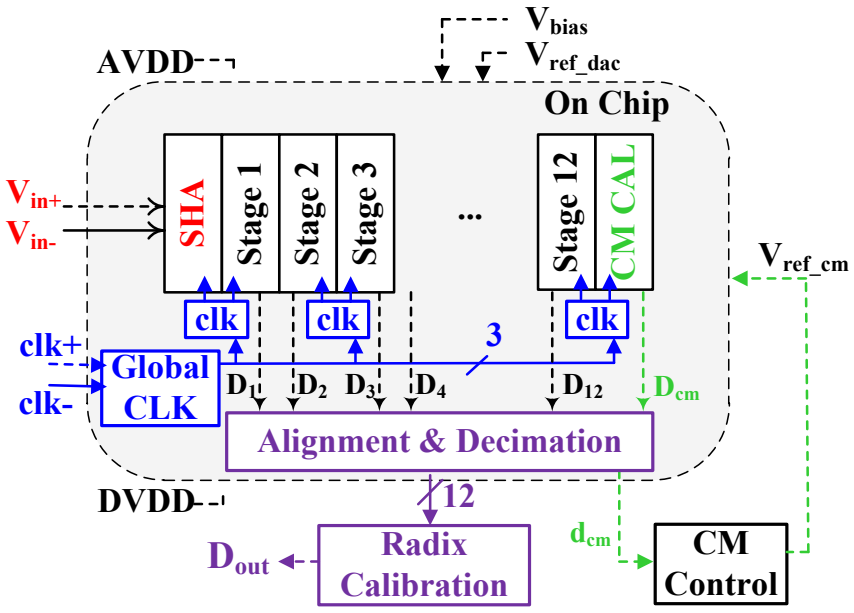


Figure 5.2 – The complete prototype ADC's architecture

With this clocking scheme, the bias and adjust phases occupy a total of ~15…20 % of the conversion cycle, while the remaining portion (~10…15 %) is for

clocking overheads (i.e. finite rise/fall times and non-overlap times between clock phases). With a settling budget of 650…700 ps for amplification, the interstage gain is estimated upon layout extraction to be $G \sim 1.55…1.65$, as discussed in Chapter 4. As a result, to achieve a quantization resolution of $B = 7…8$ bits, the ADC requires a minimum number of $N = log(2^B) / log(G) \sim 12$ stages. To implement our comparator look-ahead scheme (more in Section 5.4), a front-end SHA is needed. With also a back-end redundant stage for CM calibration, the complete ADC consists of 14 stages as shown in Figure 5.2. No stage scaling was implemented in this experimental converter.

The digital outputs $D_1...D_{12}$ from the stages, due to their half-cycle offset, are collected, aligned, and decimated on-chip before getting routed off-chip for radix calibration [29]. Decimation by a large factor (81x in this work) is required for this converter to bring its high data output rate of 500 MS/s down to the low MHz range, making it possible for CMOS drivers to drive the data off-chip. As will be discussed in Section 5.4.3, the last bit d_{cm} is used in an off-chip control loop to simultaneously calibrate the common modes of all 14 stages, similar to the work in [30].

5.2. Clock Design and Data Acquisition Flow

Referring to Figure 5.1, the ADC requires a non-uniform duty cycle, non-overlapping four phase clocking scheme. To implement this, the ADC uses off-chip fully differential clock signals (i.e. clk+ and clk- in Figure 5.2 and Figure 5.3) that are buffered through current mode logic (CML) buffers and converted to CMOS on chip (see Figure 5.3). This CMOS clock signal is then skewed by a programmable inverter chain to generate two unequal non-overlapping phases by the global clock. These three clock signals are then distributed to seven local clock generators (which are shared by the 14 stages on a pair-wise basis to minimize power) to generate all the necessary phases for stage operation.

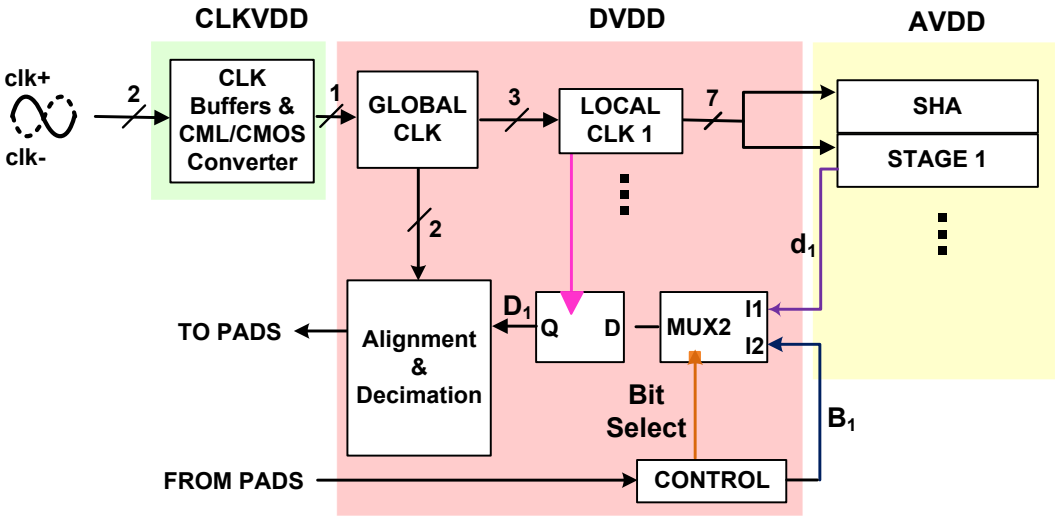


Figure 5.3 - The data acquisition flow

As shown in Figure 5.3, three different power supplies (CLKVDD, DVDD, and AVDD) are used on chip to separate the clock buffers, the digital section, and the analog section from one another. Using a programmable control block, the digital bits $D_1 \ldots D_{12}$ into the alignment and decimation block can be multiplexed between either the stage output bits ($d_1 \ldots d_{12}$) or the programmable static bits ($B_1 \ldots B_{12}$). With this data acquisition scheme, the debugging effort is alleviated during chip testing, since errors from the acquisition network (flip-flops, alignment and decimation block, CMOS drivers, etc.) can be separated from the ADC's stage errors.

5.3. Front-End Sample-and-Hold (SHA)

In our prototype ADC, the dynamic amplifier discussed in Chapter 3 and 4 was used to implement the front-end SHA. Figure 5.4 shows its complete circuit implementation. During Φ_1, the amplifier is reset by grounding its Gate and shorting its Drain and Source to V_{bias}, which is chosen so that, during the input sampling phase Φ_2, the transistor is biased in weak inversion (i.e. $V_{gs} \sim 100 \ldots 150$ mV). During Φ_2, the drain and source are connected to a negative half circuit (not shown), establishing a pseudo-differential configuration. Following Φ_2, Φ_3 is an adjust phase in which the capacitor C_{cm} (pre-charged to V_{ref_cm}) is switched to the gate to fine-tune the output

common mode levels. Finally, during Φ_4, amplification occurs, and the amplifier drives the next stage's sampling capacitance.

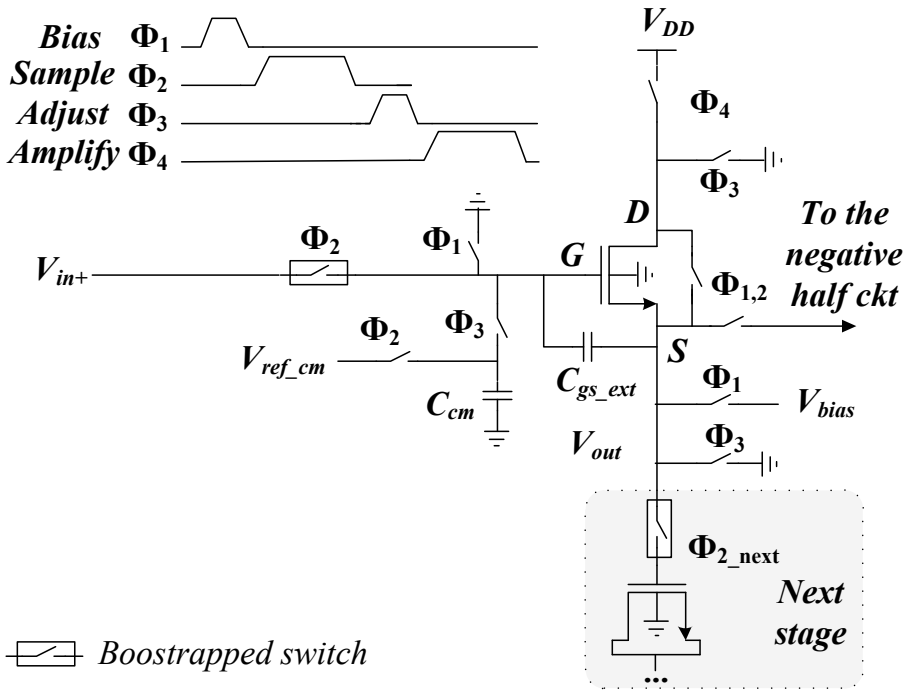


Figure 5.4 – Schematic and operation of the front-end sample-and-hold circuit

Due to the addition of the common mode capacitance C_{cm} for common mode tuning, Appendix B shows that the amplifier's ideal common-mode transfer function is modified from (3.5) to give

$$V_{oc} = \frac{C_{pg}+C_{gs}+C_{gd}}{C_{gd}+C_{pg}}V_{ic} - \frac{C_{gs}+C_{gd}}{C_{gd}+C_{pg}}V_X - \frac{C_{gs}+C_{gd}+C_{gp}}{C_{gd}+C_{pg}}V_T + \frac{C_{gd}}{C_{gd}+C_{pg}}V_{DD} + \frac{C_{cm}}{C_{gd}+C_{pg}}V_{ref\_cm} \quad (5.1)$$

With V_{ref_cm} as a tuning knob, the amplifier common mode output V_{oc} can be adjusted to its desired level to optimize the amplifier's performance. From linearity considerations in Chapter 3, V_{oc} is chosen through simulation and design iterations to be about $\frac{1}{2}V_{dd}$ to achieve a maximum output swing of ~600 mVpp while maintaining a low variation (~3 %) in sampling capacitance C_s. Overdriving each stage beyond this range generally increases C_s variation and pushes the amplifier closer to its triode

region and enhances the amplifier's distortion. To keep the amplifier's overall linearity above the 8-b level, the output of the single-ended amplifier is bounded by design by

$$450mV \approx V_{oc} - \frac{1}{2}V_{od\_max} \leq V_{out} \leq V_{oc} + \frac{1}{2}V_{od\_max} \approx 750mV \tag{5.2}$$

5.4. STAGE DESIGN

Besides the dynamic amplifier, a comparator and single-bit capacitive DAC were added to realize a complete pipelined stage, as shown in Figure 5.5. In this figure, the amplifier of Figure 5.4 is abbreviated as a triangular symbol. The comparator acquires the input signal (together with the amplifier) during Φ_{2L} and defines the polarity of charge addition/subtraction via C_{dac} when Φ_{latch} goes high. To enable high sampling rates, the input sampling switches of every stage following the SHA are enabled early (i.e. long Φ_{2L}, see timing diagram in Figure 5.5), before the previous stage enters its amplification phase (Φ_{4_prev}).

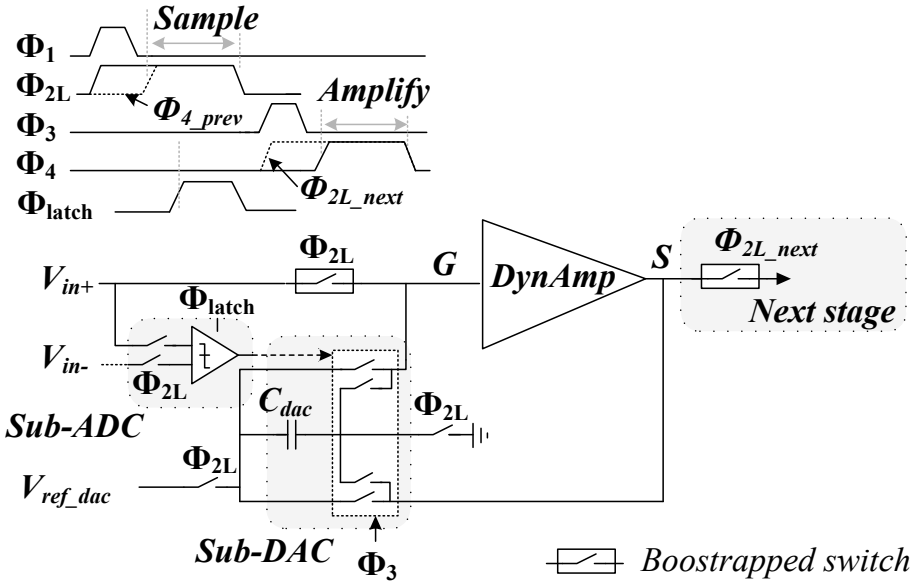


Figure 5.5 – Schematic and operation of a complete ADC stage

Similarly to (5.1), Appendix B shows that the addition of the DAC capacitance C_{dac} leads to the following revised ideal transfer function of the stage's differential mode, which is given by

$$V_{od} = \frac{C_{gs} + C_{gd} + C_{pg}}{C_{gd} + C_{pg}} V_{id} - \frac{C_{dac}}{C_{gd} + C_{pg}} V_{ref\_dac} \quad (5.3)$$

5.4.1. COMPARATOR LOOK-AHEAD SCHEME

Referring to Figure 5.5, the total cycle time (T_s) of the ADC is limited by the signal propagation delay from the stage's input through the comparator, the DAC, and the amplifier to the stage's output. Among these blocks, the amplifier's settling time and the comparator's delay are most significant. With incomplete settling, the amplifier's settling time is reduced to ~700 ps or twice as large as the comparator's delay (which is ~300…400 ps for a target metastability rate of 10^{-11}). Neglecting other timing overheads (finite rise and fall times, etc.), these delays lead to a minimum T_s of 2.2 ns, making it impossible to implement a 500 MS/s converter. To solve this problem, we used a comparator look-ahead scheme that exploits the monotonicity of the stage input signal and the rapid exponential regeneration of a latch comparator to significantly reduce the cycle time. This is conceptually illustrated in Figure 5.6.

Figure 5.6 shows stage k's amplifier output $V_{od}(k)$ (the top plot) and stage k+1's comparator output D_{k+1} for two cases of implementation (denoted as Typical and Early in the bottom two plots). In the first case, which is the typical implementation of a pipelined ADC, the comparator of stage k+1 is activated only after the output of stage k settles, as illustrated by the middle curve (i.e. Typical D_{k+1}) which rises after t_2. The total cycle time in this case, as mentioned previously, is roughly twice the sum of the amplifier's settling time T_{amp} and the comparator's delay T_{compl}.

In the second case, the comparator is activated as soon as the polarity of its input signal becomes available (after a short delay Δt), as illustrated by the bottom plot (Early D_{k+1}) of Figure 5.6. Since the input into each stage after the SHA is always a monotonic signal, the comparator's final output, which is set only by the input signal's polarity, does not depend on the time of activation. As a result, early activation allows the settling time of stage k+1's comparator to be completely absorbed into that of the preceding stage's amplifier (stage k), eliminating this timing overhead from the total cycle time as a result. The drawback of this technique, however, is a reduced input

amplitude seen at time t_1 by the comparator that causes its settling time to increase (i.e. $T_{comp2} > T_{comp1}$). Nevertheless, the extra time available for exponential regeneration (which is equal to T_{amp} - Δt - T_{comp2}) compensates for this input reduction and effectively leads to a similar metastability rate. With this look-ahead scheme, T_s is now limited only by the settling time of the amplifier ($T_s \approx 2T_{amp} \approx 1.4$ ns), making it possible to achieve a conversion rate of 500 MS/s.

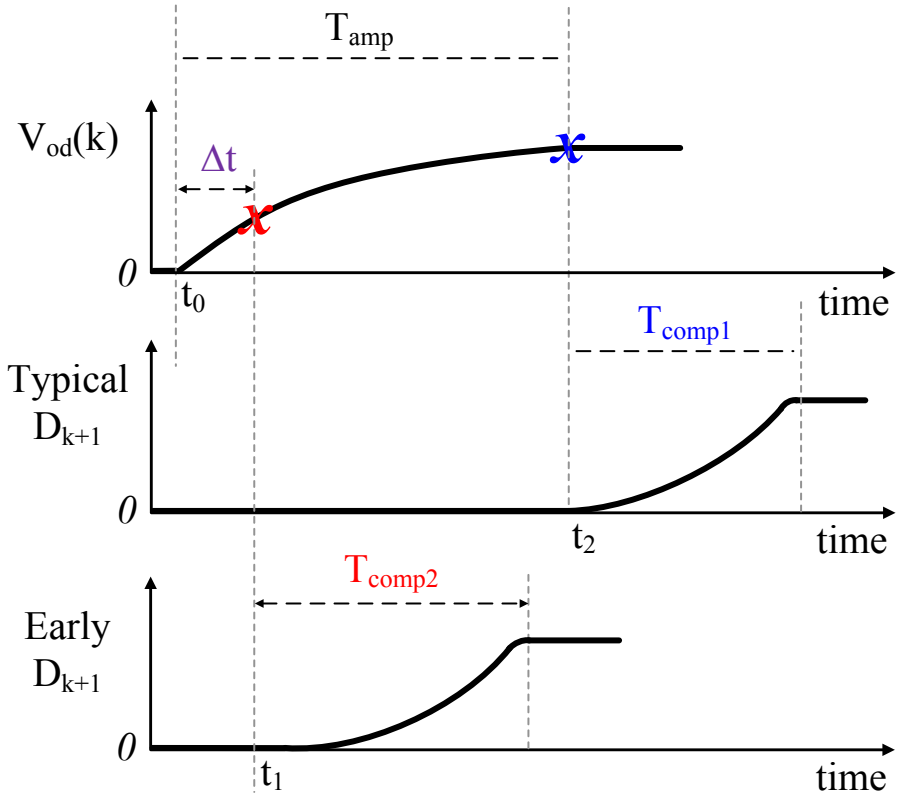


Figure 5.6 – Comparator look-ahead scheme

5.4.2. COMPARATOR DESIGN

The structure of the comparator is a purely dynamic and self-calibrating latch with a pre-amp, similar to the work in [31] as shown in Figure 5.7. The operation of the comparator is as follows. When Φ_{latch} is low, the comparator resets. M_0 is turned off, and the M_3 transistors are on. This resets D_{im} and D_{ip} to low, turning on M5 and M6 and resetting V_{op} and V_{om} to high. When Φ_{latch} goes high, the comparator begins to

regenerate the input signal. M3 is turned off, and M0 is turned on. D_{ip} and D_{im} now rise at different rates set by the input signals V_{ip} and V_{im} at the gates of the M1 transistors. As D_{ip} and D_{im} continue to rise, M5 and M8 are turned off, and M4 is turned on. V_{op} and V_{om} begin to fall (at different rates) from V_{DD} and eventually diverge (and settle) to their final values due to the positive feedback action of the cross-coupled inverters of M6 and M7.

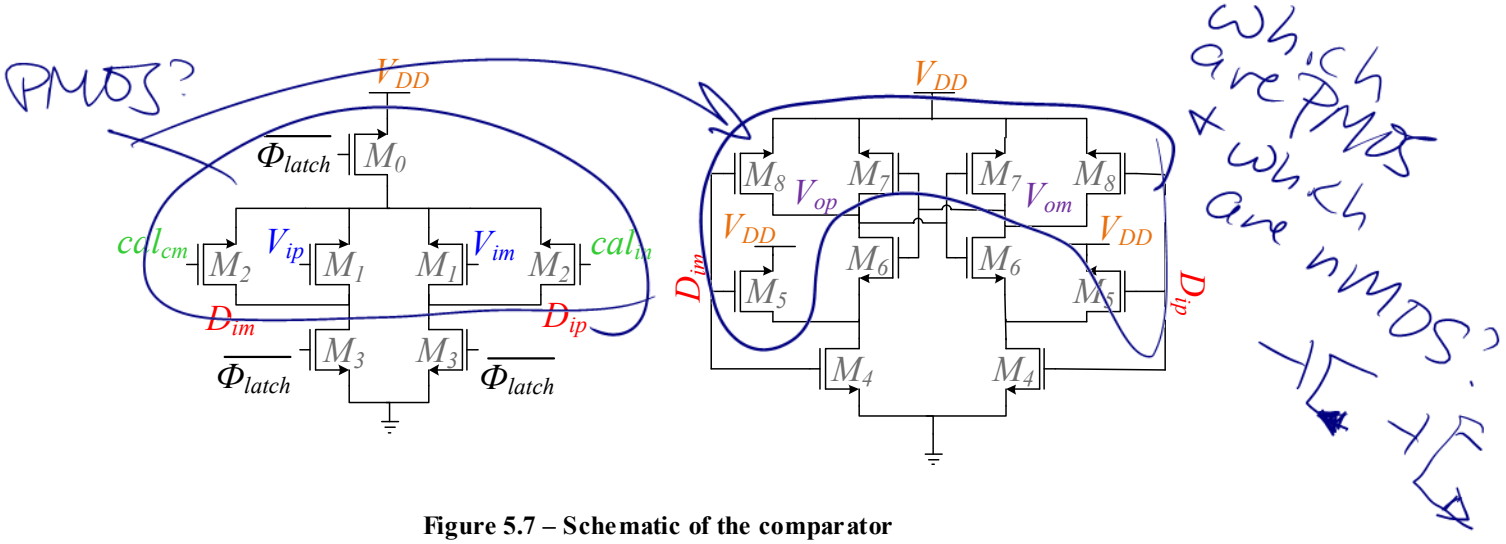


Figure 5.7 – Schematic of the comparator

Due to the look-ahead scheme implemented in this work, Φ_{latch} goes high when both V_{ip} and V_{im} at the comparator input are still rising to their final values (from a zero potential). Thus, if the M1 transistors are NMOS, such a low input common mode would result in a low g_m and hence a slower response time of the input pre-Amp. To ensure a fast response time, the input stage is therefore implemented with a PMOS pair as shown in Figure 5.7.

In this scheme, the M2 transistors are used for offset calibration in a bang-bang control loop similar to [30, 31]. Specifically, since the comparator is only needed during Φ_1 and Φ_2 to latch the input signal (see Figure 5.5), offset calibration is performed during Φ_3 and Φ_4 in an identical fashion but with a zero input signal. This technique leads to a near minimum size, low offset, high speed, and low power comparator suitable for our prototype ADC.

5.4.3. COMMON MODE CALIBRATION

To calibrate for the common mode of all the stages, the output common mode of the last ADC stage (stage 12) is sampled by the CM Cal stage and compared against the desired voltage V_{cm_des}, as shown in Figure 5.8. The resulting bit of this comparison, d_{cm}, indicates if a higher or lower voltage V_{ref_cm} is needed to fine tune the common mode output [see equation (5.4)]. V_{ref_cm} is thus manually adjusted off-chip until d_{cm} dithers randomly between 0 and 1.

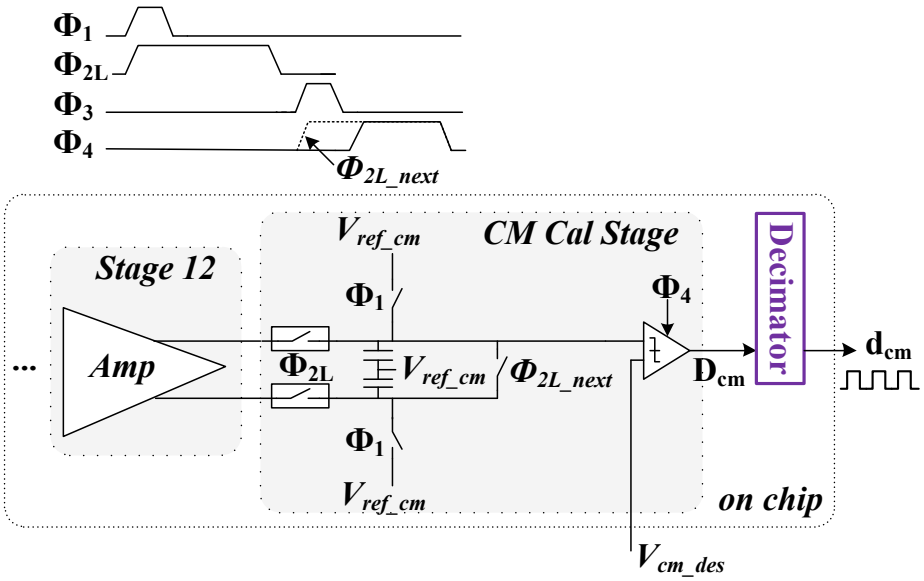


Figure 5.8 - The CM control loop

5.5. LAYOUT STRATEGY

To optimize the ADC's performance at high speed, we designed the ADC's layout to be similar to the block diagram of Figure 5.2, with an emphasis on minimizing not only the overall dimension of the critical blocks (comparators, amplifiers, etc.) but also the ADC's dimension along the signal path (the path from the SHA to the CM Cal stage). This not only reduces the parasitic loading in the clock distribution network but also minimizes the analog signal propagation path from stage to stage. In this work, the layout dimension of each stage is approximately 50 µm x

19.5 µm (see Figure 5.9), with the shorter geometry being in the direction of signal propagation.

Regarding the structure of each ADC stage, the comparator network was laid out adjacently to the amplifier of the ADC stage. Large (poly) capacitors were used for comparator offset calibration (see Section 5.4.2). The layout of the DAC, CM control blocks, and amplifiers of the positive half mirror that of the negative half. Additional effort was made with the layout of the amplifier's main transistor to minimize the various parasitic capacitances that could potentially reduce the voltage gain (see Section 3.2). In particular, to minimize gate parasitic capacitances and maximize C_{gs_ext} for a given layout area, the layout of C_{gs_ext} was designed such that the source terminals completely surround the gate terminals, as shown in Figure 5.10. Additionally, the number of drain contacts was halved, and body contacts (which are shared by the two circuit halves) were pulled slightly away from the gates to further reduce their contribution to the gate parasitic capacitances (see Section 3.4). The combination of these layout efforts led to an estimated gain of ~1.55…1.65 (upon parasitic extraction) for an amplifier settling time of ~700 ps.

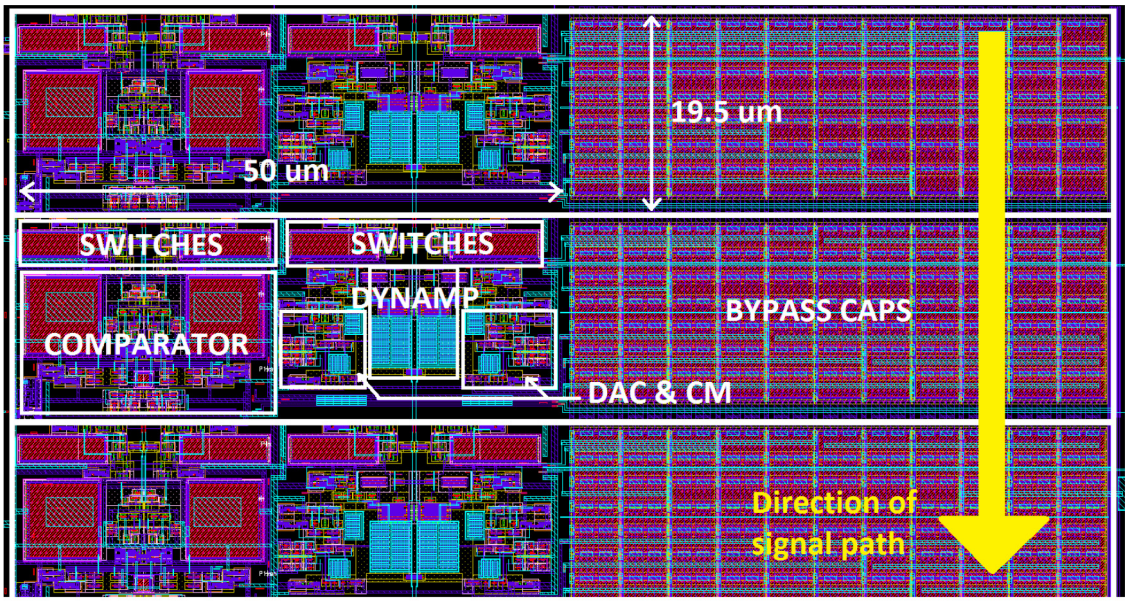


Figure 5.9 - Stage's layout

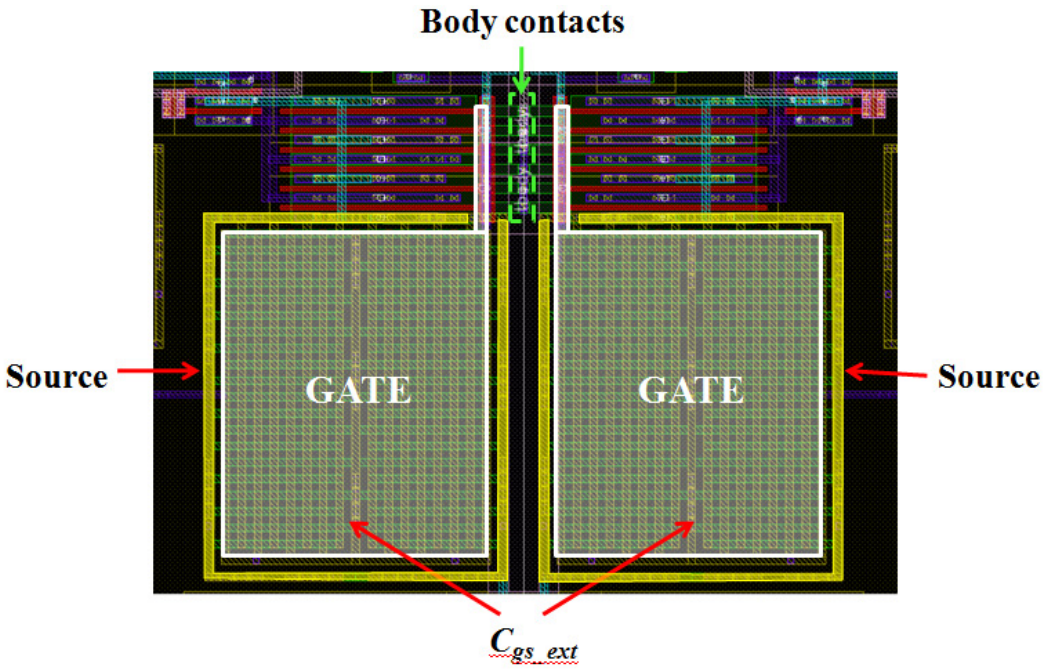


Figure 5.10 - Amplifier's layout

5.6. CHAPTER SUMMARY

This chapter discussed the implementation details of our prototype ADC and showed how the incomplete settling technique described in Chapter 4 can be combined with a comparator look-ahead scheme to achieve a conversion rate of 500 MS/s. A layout strategy to reduce parasitic capacitances was also discussed. The next chapter will present the measurement results of our prototype ADC.

CHAPTER 6. MEASUREMENT RESULTS

The prototype ADC was fabricated by STMicroelectronics in a 65-nm SOI CMOS technology. Figure 6.1 shows the active area of the ADC which consists of 14 stages and occupies an area of 270 µm x 70 µm. Figure 6.2 shows the photograph of the entire die, including the bond wires. The chip occupies a total area of 1mm x 1mm and was packaged in a 48-pin QFN (quad flat no lead) package.

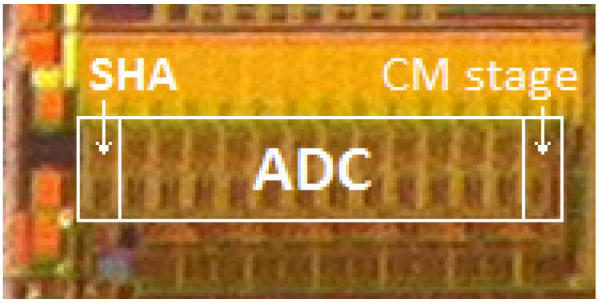


Figure 6.1 - Die photograph of prototype ADC

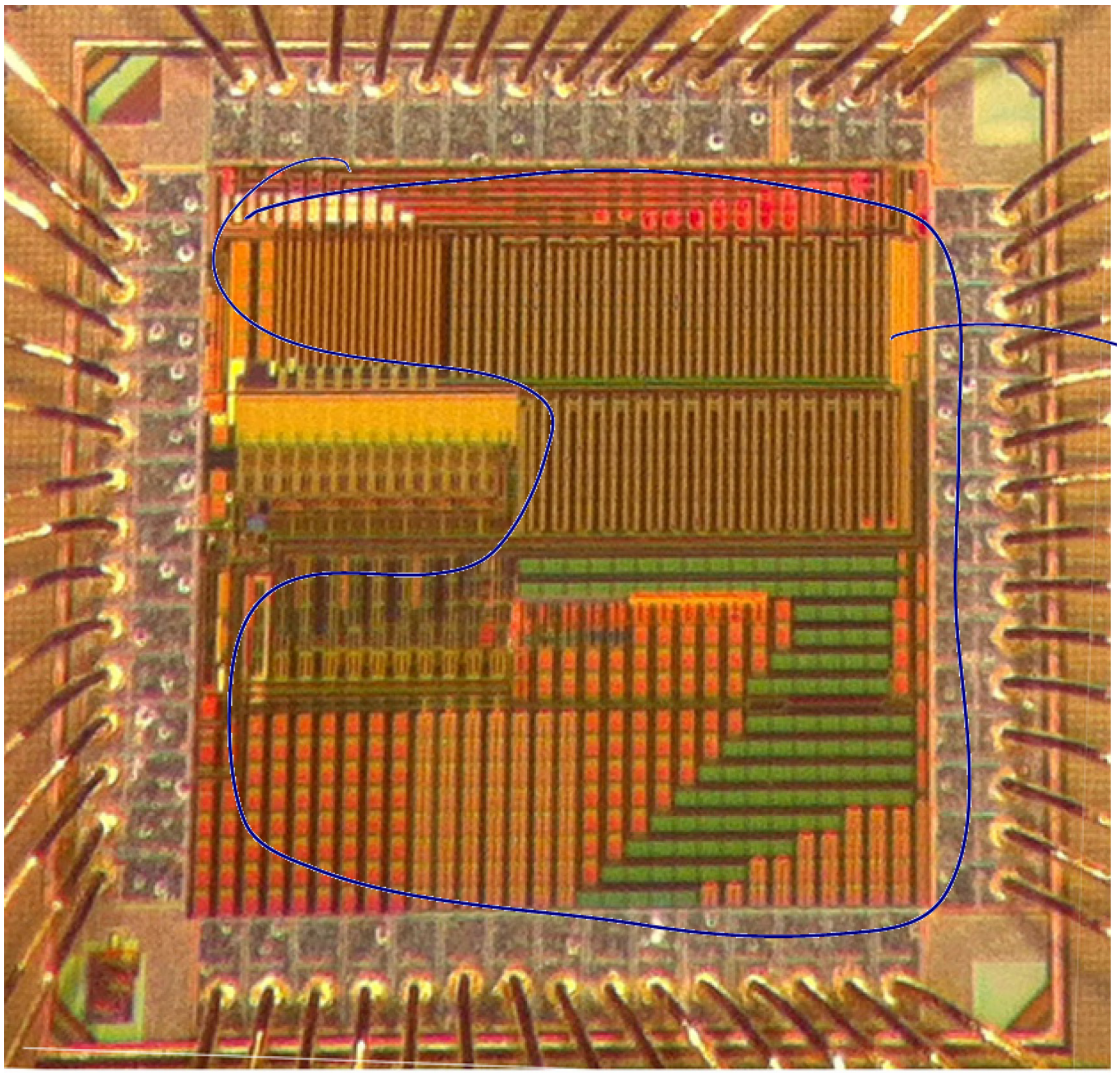


Figure 6.2 - Die photograph (including bond wires)

6.1. TEST SETUP

Figure 6.3 shows the measurement setup. To test our prototype ADC, we used a custom designed two-layer evaluation board. The chip's package was attached to the board by a tension socket during testing. The voltage references required to operate the ADC are generated on board by op-amps and set by resistive ladders. Both the clock and input sinusoidal signals are passed onto the chip via transformers. All relevant clock sources are phase locked via a common 10-MHz reference signal. The digital outputs of the chip are converted from CMOS to LVDS (low-voltage-differential signaling) on board and captured by the NI (National Instruments) 6562

IO card. Labview was used to program the on-chip shift registers and process the digital output data. The final data analysis was performed in Matlab. For radix calibration, we used a low frequency input signal (~25 MHz) to measure the stage radix coefficients which were used in all subsequent measurements.

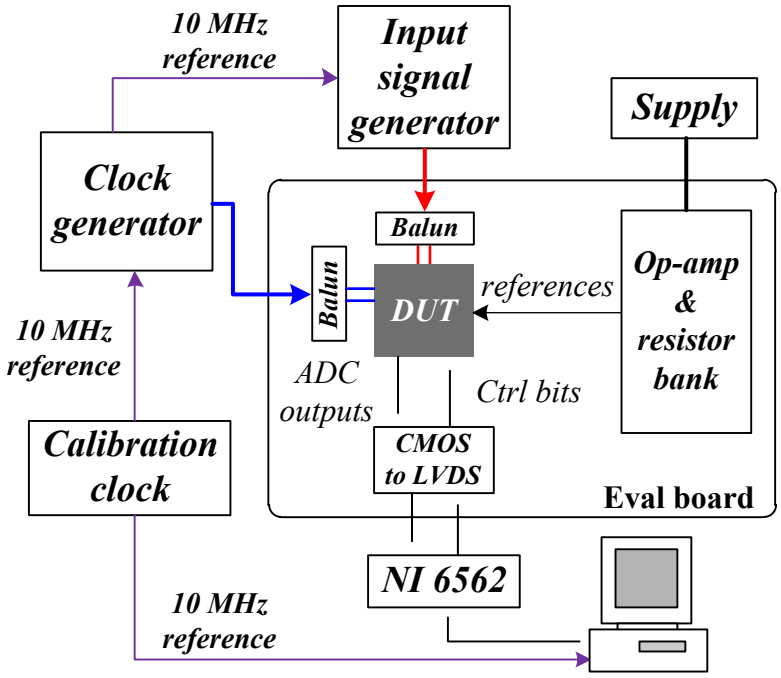


Figure 6.3 - Test setup

6.2. MEASUREMENT RESULTS

6.2.1. LINEARITY

Figure 6.4 shows the ADC's measured (decimated) output spectrum with f_s = 502.2 MS/s and a 600-mVpp input signal at Nyquist frequency (f_{in} ~ 251 MHz). This spectrum shows a signal-to-noise-and-distortion ratio (SNDR) of ~41.5 dB and a spurious free dynamic range (SFDR) of 53.6 dB, showing that the ADC is linear up to 8…9 bits as expected from the linearity analysis in Chapter 3. Furthermore, from Figure 6.5, it is apparent that the performance of the ADC remains relatively flat for a wide input frequency range up to Nyquist. Beyond the Nyquist frequency, however, the

ADC performance begins to degrade, primarily due to an increase in distortion indicated by a reduced SFDR.

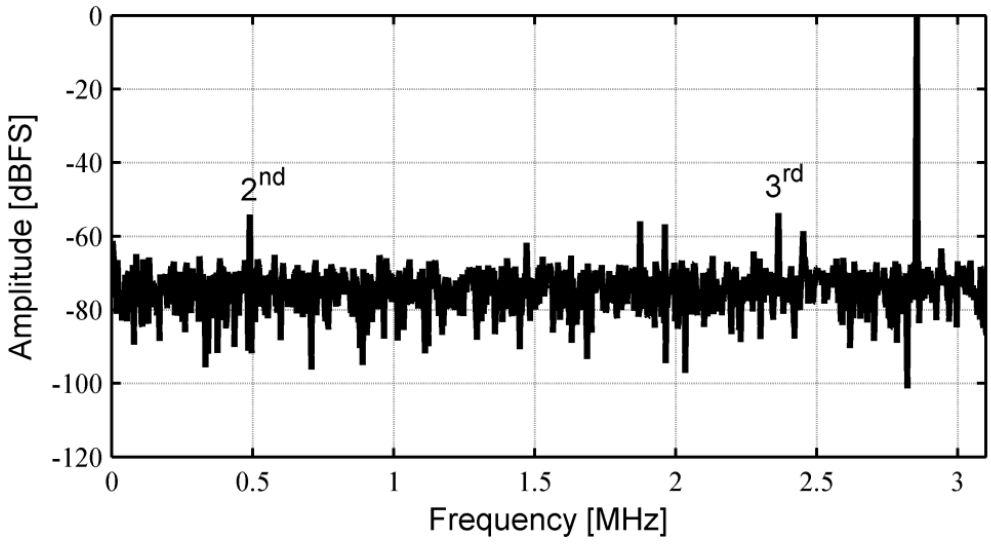


Figure 6.4 - Measured output spectrum (2048 point FFT) after decimation by 81 with f_s = 502.2 MS/s and Nyquist input (f_{in} ~ 251 MHz, SNDR = 41.5 dB, and SFDR = 53.6 dB)

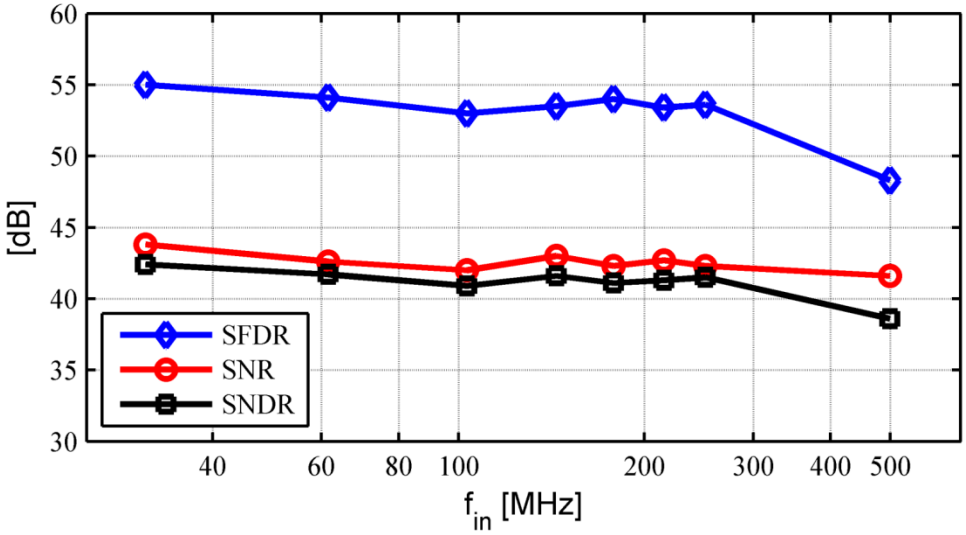


Figure 6.5 - Measured performance vs. f_{in}

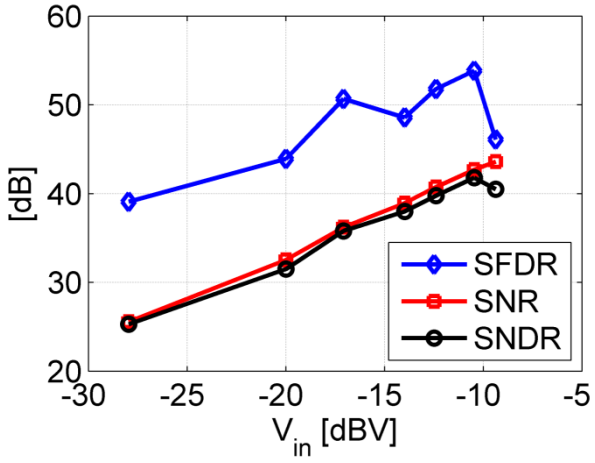


Figure 6.6 - Measured performance vs. input amplitude (f_{in} ~ 25 MHz)

Figure 6.6 shows the ADC performance vs. input amplitude, illustrating a maximum input swing of approximately -10.5 dBV (or ~600 mVpp). Both Figure 6.5 and Figure 6.6 suggest that the ADC's performance is limited not by the dynamic amplifier's distortion but primarily by noise, which consists of thermal noise, differential nonlinearity (DNL) noise, inter-stage jitter induced noise (due to incomplete settling, see Chapter 4), and quantization noise. A histogram measurement in which the ADC input was held constant to observe the variation in the output bits showed that the overall noise was about 1.8 mV-rms (or ~0.9 LSB-rms), corresponding to an overall peak SNR of approximately 41.4 dB (which agrees with the measured spectrum of Figure 6.5).

6.2.2. DNL AND INL

From Figure 6.7, the measured DNL and INL (differential and integral nonlinearity) are +0.61/-0.65 LSB and +1.07/-1.15 LSB, respectively. The DNL measurement shows a total of 295 levels with no missing codes, corresponding to an average amplifier gain of $295^{1/12}$ ~1.6. The INL is primarily limited by the nonlinearity of the front-end SHA, the stages' dynamic amplifiers as well as their incomplete transient settling (see Chapter 4).

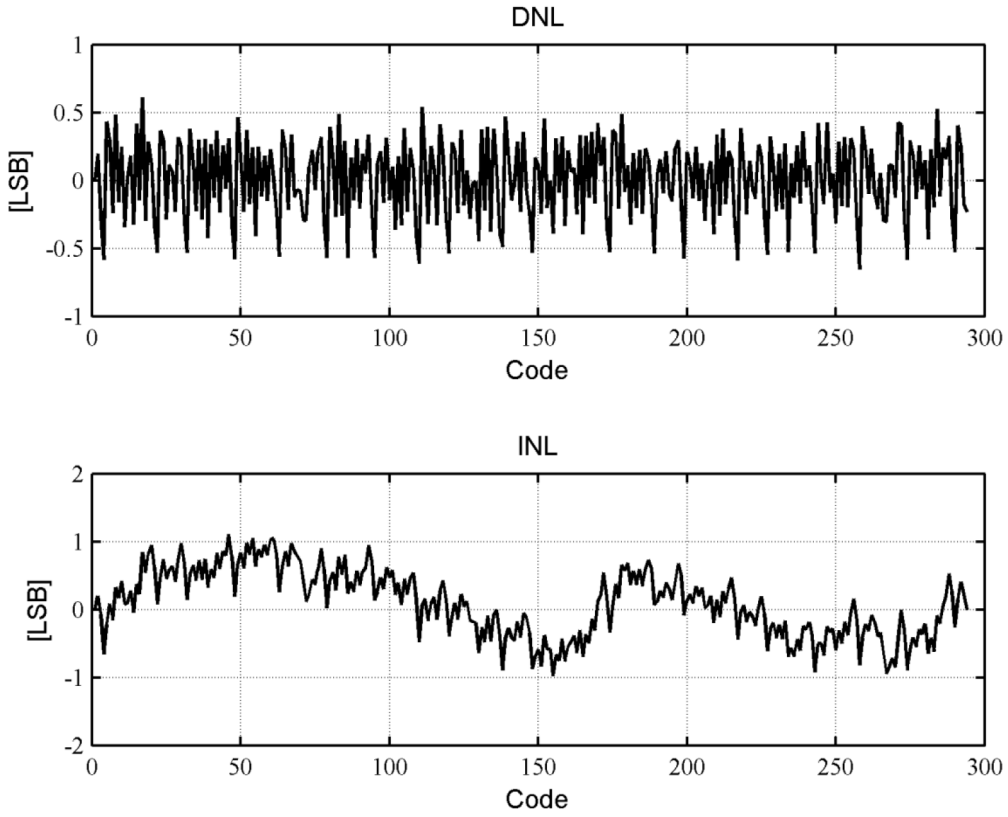


Figure 6.7 - Measured DNL (+0.61/-0.65 LSB) and INL (+1.1/-0.97 LSB) at f_s = 500 MS/s

6.2.3. INPUT COMMON MODE SENSITIVITY

To observe the ADC's sensitivity to the input CM, the performance of the ADC is recorded as the input CM is varied, as shown in Figure 6.8. The left plot shows the performance of the ADC with fixed calibration coefficients obtained at V_{ic} = 0.6 V. The right plot shows the performance with re-calibration done at every V_{ic} value. Both plots show a minor variation in recorded performance, demonstrating the ADC's relatively good rejection to common mode fluctuation. This observation aligns with prior art's reported results in [11].

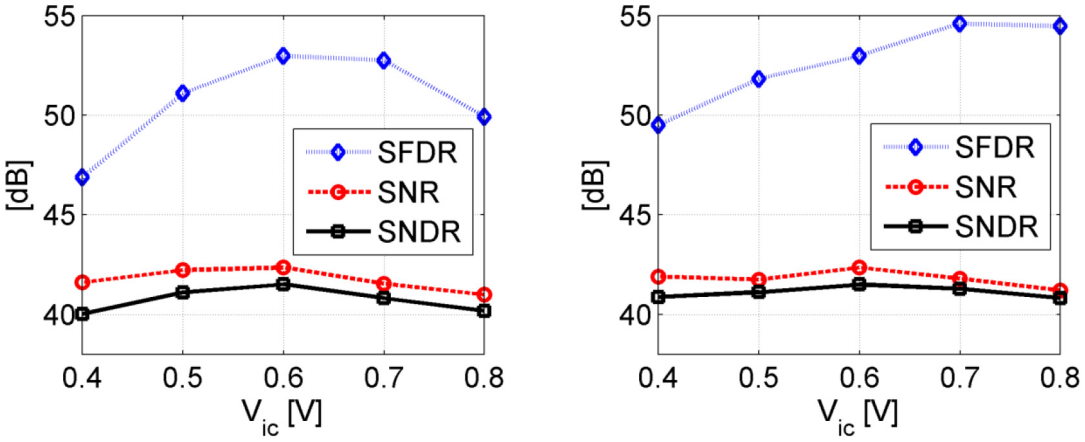


Figure 6.8 - ADC performance vs. input CM with calibration only at V_{ic} = 0.6 V (left plot) and with re-calibration at every V_{ic} value (right plot)

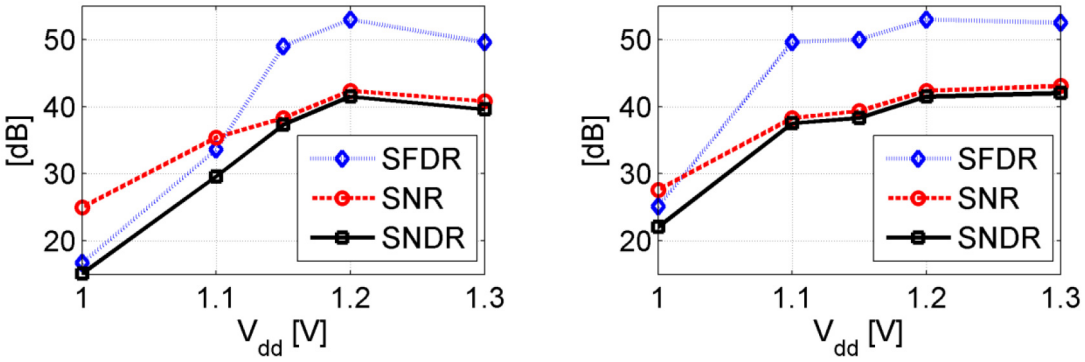


Figure 6.9 - ADC performance vs. V_{dd} with calibration only at V_{dd} = 1.2 V (left plot) and with re-calibration at every V_{dd} value (right plot)

6.2.4. SUPPLY SENSITIVITY

As can be seen in Figure 6.9, the performance of the ADC is sensitive to supply variation. If the radix coefficients (obtained at V_{DD} = 1.2 V) are kept constant, the degradation in SNDR is approximately -12 dB per 100 mV of supply reduction. With re-calibration performed at each supply value, the degradation in SNDR is about -5 dB at V_{DD} = 1.1 V. Below this value, however, the performance degrades dramatically even with re-calibration. This is expected, because the amplifier draws its current directly from the supply, and, due to incomplete settling, its gain at the end of amplification strongly depends on the device's transconductance g_m, which is

proportional to the supply level (see Chapter 4). A reduced supply thus leads to a smaller gain and higher quantization noise level that limits the ADC's performance.

6.3. Performance Summary

In summary, the experimental converter advances the state-of-the-art (see Figure 1.2) by achieving an attractive combination of measured performance which consists of a high conversion rate (500 MS/s), low conversion energy (98 fJ/conv-step), and low input capacitance (55 fF). The prototype ADC is 8-bit linear and dissipates a total power of 5.1 mW. The majority of this power (75 %) is consumed by the digital circuitry (comparators and clocks), while the remaining (25 %) is dissipated by the analog amplifiers. Furthermore, these measurement results agree well with estimated values from our analysis of Chapter 3 and 4. Table 6.1 summarizes the achieved performance. In the next chapter, we will discuss future design directions to further improve this architecture.

Table 6.1 - ADC's performance summary

Parameter	Value	
Process	65-nm SOI CMOS	
Active area	0.02 mm^2	
Supply V_{dd}	1.2 V	
Input capacitance	55 fF	
Full Scale Range	600 mVpp	
Resolution	8.2 b (295 levels)	
Sampling Frequency	502.2 MS/s	
	$f_{in} \approx 25$ MHz	$f_{in} \approx 251$ MHz
SNDR	42 dB	41.5 dB
SFDR	55 dB	53.6 dB
DNL	+0.61/-0.65	
INL	+1.1/-0.97	
Power		
Analog (Amplifiers)	1.2 mW	
Digital (Comparators, Clocks)	3.9 mW	
Total	5.1 mW	
FOM	98 fJ/conv-step	

CHAPTER 7. CONCLUSION & FUTURE WORK

7.1. CONCLUSION

This work has demonstrated the analysis and detailed implementation of an 8-bit pipelined converter that advances the state-of-the-art by simultaneously achieving a high sampling rate (500 MS/s), low power (5.1 mW), and low input capacitance (55 fF). The achieved FOM is 98 fJ/conv-step, which places this converter as one of the most energy efficient converters above 150 MS/s and is 3x better than competing pipelined architectures within the same frequency range. Key techniques that were employed in this work consist of dynamic reside amplification, incomplete settling, and a comparator look-ahead scheme that eliminates the timing overhead of the sub-converters to maximize the ADC's speed. At this point, there exist interesting opportunities to further improve this architecture both in performance and robustness. Although the proposed technique was implemented with a 65-nm SOI CMOS process, modification to the technique was also discussed in an attempt to make it more suitable

for a Bulk CMOS implementation. It thus remains interesting to see within the foreseeable future if this architecture will gain adoption by industry, not only in A/D conversion but also in other low-power applications.

7.2. Future Work

7.2.1. Power Saving with Stage Scaling

An interesting aspect that can be explored to further improve the power efficiency is to employ stage scaling. Since all the stages in this experimental converter are of identical size, much of the power is wasted because all the stages after the first stage burn additional power to drive large capacitances that do not contribute much to the ADC's overall noise. For a high speed converter such as this, a better approach is to scale down the stages by a factor approximately equal to the stage's gain [15]. The resulting power saving that comes with stage scaling can in turn be invested in reducing the first stage's noise (by making its sampling cap larger), thereby improving the overall SNR of the converter.

7.2.2. Power Saving with an Innovative Layout Design

The layout of this experimental converter, as discussed in Section 5.5, is not the most power efficient solution. In this work, since the stages were laid out linearly. For a brute force approach in which all the necessary ADC phases are generated by a single global clock and distributed to the stages, the cost to drive such a large clock bus is prohibitively expensive (see Figure 7.1a). Thus, it was found through simulation and parasitic extraction that breaking the clock generation into two steps (global and local) to minimize the number of long wires was a far more power efficient solution (see Figure 7.1b). Nevertheless, this distribution network still consists of large buffers and long wires and, as a result, consumes approximately half of the total digital power. Taking into account stage scaling, an even better approach would be to layout the stages in a circular fashion to minimize the wiring and buffering power (see Figure 7.1c). With a reduced wiring overhead, the local clocks can also be eliminated, and a

single global clock to generate all the phases would be sufficient. One may also explore opportunities in innovative and low power clocking schemes to further optimize the converter's efficiency, since, currently, 75 % of the total power is dissipated by the clocks and comparators.

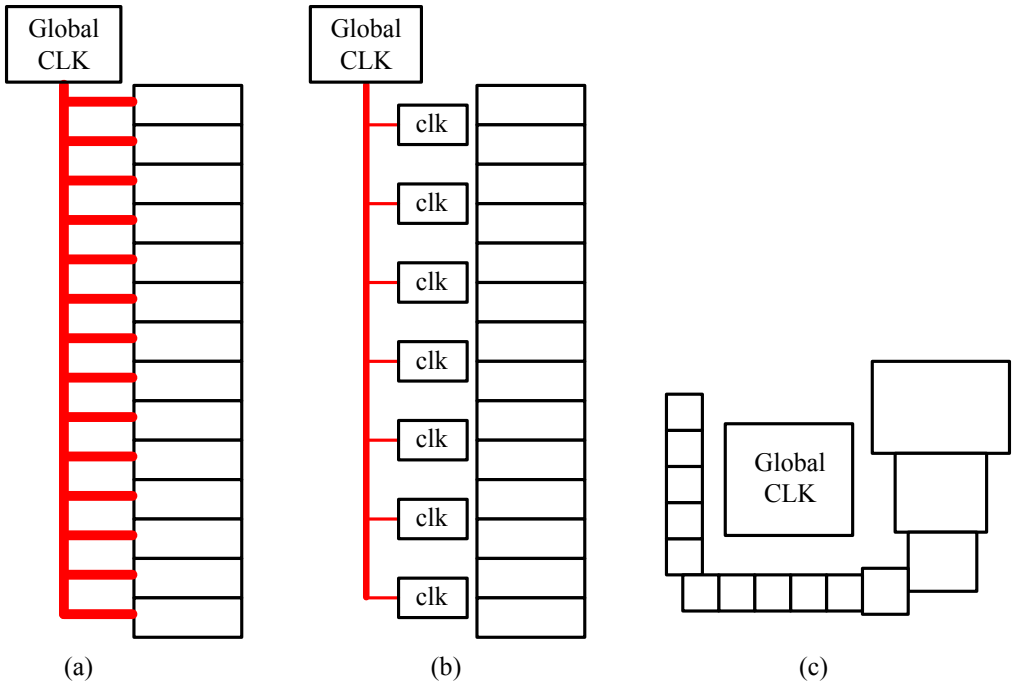


Figure 7.1 - Layout options that are (a) wasteful (b) more efficient (c) most efficient

7.2.3. On-chip Continuous Calibration of the Amplifier's CM

For this experimental converter, the CM calibration was performed off-chip using the last bit d_{cm} as an indication of the CM level (see Section 5.4.3). In future implementations, this adjustment loop can be added on chip with a small power overhead. ~~One possible~~ way of implementing this loop is to employ the same principle as the band-bang control loop used to calibrate for comparator offsets (see Section 5.4.2). Such a loop can run continuously in the background to keep the ADC always at its desired bias point.

ok, if you are going to reference it twice, please explain the method in the text.

7.2.4. CALIBRATION FOR TEMPERATURE VARIATION

In this work, no measurement of the converter's temperature sensitivity was carried out. Nevertheless, prior research has found that the converter's performance stays relatively constant if re-calibration (of the radix coefficients) is performed along with the temperature variation [30]. To examine the dynamic amplifier's sensitivity to temperature changes, a temperature sweep was done in simulation. As shown in Figure 7.2, simulation shows that both the amplifier's gain G and output CM V_{ocm} vary monotonically with temperature. Here, it is crucial to note that, although there is no direct way to continuously measure G on-chip, the output CM which tracks temperature changes can continuously be monitored using the CM detection loop mentioned previously (see Sections 5.4.3 and 7.2.3). The knowledge of the relative CM change (from its desired point) can then be used to infer information about the relative temperature deviation. From this point, the radix coefficients can (in principle) also be continuously re-adjusted in the background using pre-programmed lookup tables or via interpolation. Alternatively, the CM information can also be encoded on-chip and appended to the ADC's output bits for off-chip post processing.

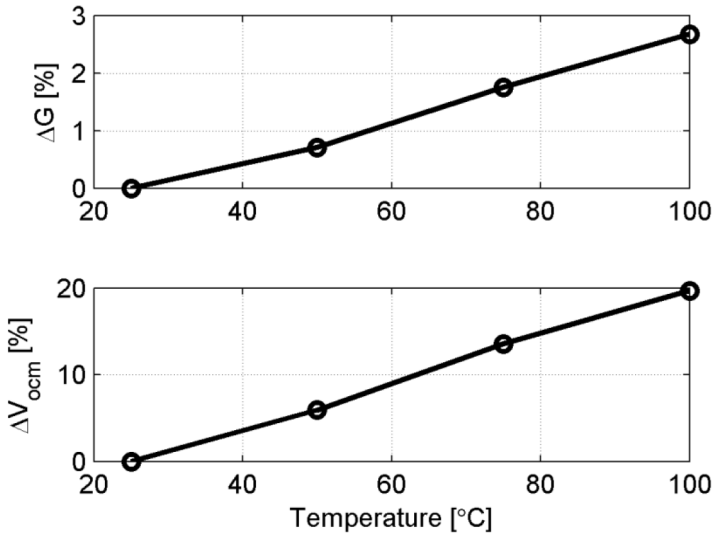


Figure 7.2 - Variation of gain and output CM with temperature

7.2.5. OTHER AMPLIFIER IDEAS

The knowledge that we gained from this work can be further leveraged by experimenting with other interesting charge-based amplification topologies. The reason for this is because these amplifier topologies share very similar characteristics as the amplifier discussed in this work. For instance, the "capacitive coupled amplifier" introduced in 1979 [32] for filter and integrator applications and the "charge transfer amplifier" used for comparator offset cancelation in Flash ADC [33] share essentially identical amplification concept with minor differences in implementation details. As a result, the analyses and results found in this work are also applicable to these topologies.

Another interesting topic that can also be explored with dynamic amplification in the future is to incorporate it with a positive feedback latch during amplification, as conceptually illustrated in Figure 7.3. The sampling phase is identical to that of the dynamic amplifier implemented in this work. The amplification phase is also almost identical; except that now, a short moment Δt_1 after Φ_4 rises, a cross coupled PMOS pair is connected to the output of the amplifier. Due to the positive feedback action, this cross coupled pair provides additional amplification by regenerating the output differential voltage. To better illustrate this, Figure 7.4 shows a typical transient progression of this amplifier during amplification. During the first interval, only transistors M_a's conduct currents, and the operation of this amplifier is identical to our old dynamic amplifier. As soon as the polarity is defined, Φ_{boost} goes high, and this polarity is amplified exponentially by the M_b cross-coupled pair until Φ_{boost} shuts off. With this technique, the gain of the amplifier can be significantly larger than what can be achieved without the cross-coupled pair while maintaining a high sampling speed.

The drawback of this method is an increase in distortion caused by the M_b cross coupled pair during amplification. Simulation shows that the achievable linearity of this amplifier (without calibration) is on the order of 6 bits. Nevertheless, what is attractive about this technique is that, the M_a transistors can be made very small (~15x smaller than the transistors used in this work), because most of the amplification action comes not from M_a's but from M_b's (which are also of similar widths). As a result,

sampling linearity is greatly improved, because the M_a nonlinear moscaps are now much smaller than the linear C_{gs_ext} caps. Thus, almost none of the distortion comes from the sampling caps but from the latch amplifiers, which, in principle, can be corrected to recover the lost resolution [21]. Furthermore, this technique does not require an SOI process, making its adoption for design with conventional Bulk CMOS circuits easier.

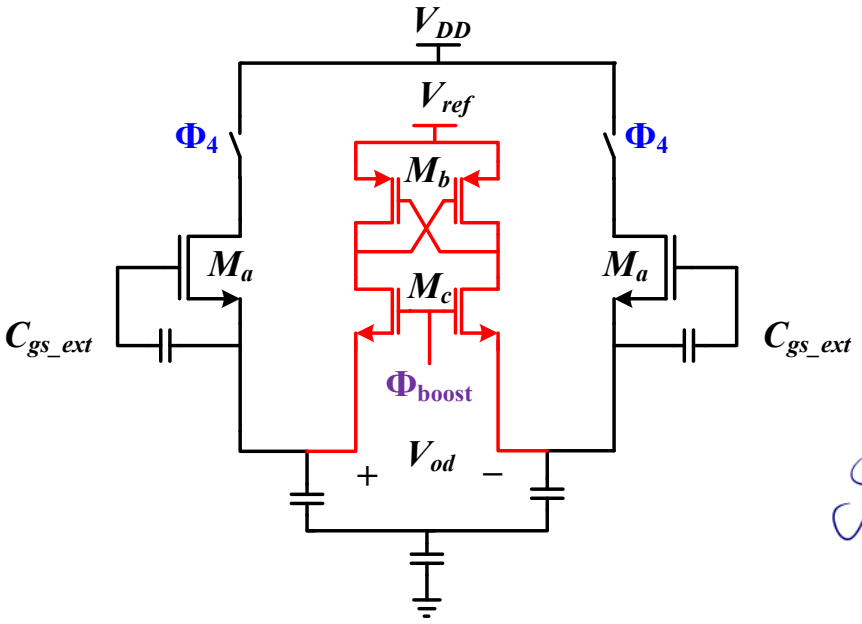


Figure 7.3 - Dynamic amplifier with a positive feedback latch

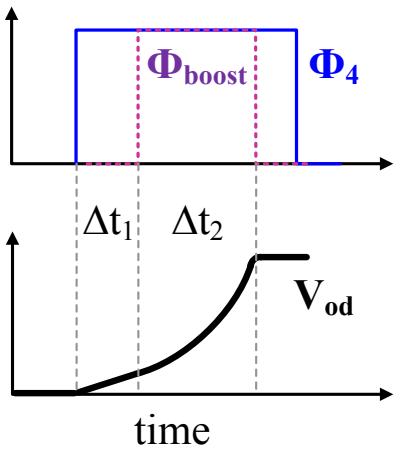


Figure 7.4 - A typical transient progression of the differential output

APPENDIX A – GAIN ANALYSIS

To transform the charge equations into more manipulable forms, we need to make a few approximations. From the equations of Chapter 3, we have

$$V_{op} = \frac{C^s_{gsdp} + C_{pg} + \frac{1}{2}\Delta C^s_{gsd}}{C_{pg} + C^a_{gd} + \frac{1}{2}\Delta C_{gd}} V_{ip} - \frac{C^a_{gsdp} + C_{pg} + \frac{1}{2}\Delta C^a_{gsd}}{C_{pg} + C^a_{gd} + \frac{1}{2}\Delta C_{gd}} V_{tp} - \frac{C^s_{gsd} + \frac{1}{2}\Delta C^s_{gsd}}{C_{pg} + C^a_{gd} + \frac{1}{2}\Delta C_{gd}} V_X + \frac{C^a_{gd} + \frac{1}{2}\Delta C_{gd}}{C_{pg} + C^a_{gd} + \frac{1}{2}\Delta C_{gd}} V_{DD} \quad \text{(A1)}$$

$$V_{om} = \frac{C^s_{gsdp} + C_{pg} - \frac{1}{2}\Delta C^s_{gsd}}{C_{pg} + C^a_{gd} - \frac{1}{2}\Delta C_{gd}} V_{ip} - \frac{C^a_{gsdp} + C_{pg} - \frac{1}{2}\Delta C^a_{gsd}}{C_{pg} + C^a_{gd} - \frac{1}{2}\Delta C_{gd}} V_{tp} - \frac{C^s_{gsd} - \frac{1}{2}\Delta C^s_{gsd}}{C_{pg} + C^a_{gd} - \frac{1}{2}\Delta C_{gd}} V_X + \frac{C^a_{gd} - \frac{1}{2}\Delta C_{gd}}{C_{pg} + C^a_{gd} - \frac{1}{2}\Delta C_{gd}} V_{DD} \quad \text{(A2)}$$

We can approximate the first term of the above equation as follows

$$First\ term \approx \left(\frac{C^s_{gsdp} + C_{pg} + \frac{1}{2}\Delta C^s_{gsd}}{C_{pg} + C^a_{gd}}\right)\left(1 - \frac{\frac{1}{2}\Delta C^a_{gd}}{C_{pg} + C^a_{gd}}\right)$$

$$\approx \frac{C^s_{gsd} + C_{pg}}{C_{pg} + C^a_{gd}}\left(1 - \frac{1}{2}\Delta X^a_{gd} + \frac{1}{2}\Delta X^s_{gsd}\right)$$

Thus, applying the same approximations to the remaining terms to give

$$V_{op} = \frac{C_{gsd}^{s} + C_{pg}}{C_{pg} + C_{gd}^{a}}\left(1 - \frac{1}{2}\Delta X_{gd}^{a} + \frac{1}{2}\Delta X_{gsd}^{s}\right)\left(V_{ip} - k_c V_{tp}\right) - \frac{C_{gsd}^{s}}{C_{pg} + C_{gd}^{a}}\left(1 - \frac{1}{2}\Delta X_{gd}^{a} + \frac{1}{2}\Delta X_{gsd}^{s}\right)V_X + \frac{C_{gd}^{a}}{C_{pg} + C_{gd}^{a}}\left(1 - \frac{1}{2}\Delta X_{gd}^{a} + \frac{1}{2}\Delta X_{gd}^{a}\right)V_{DD} \tag{A3}$$

$$V_{om} = \frac{C_{gsd}^{s} + C_{pg}}{C_{pg} + C_{gd}^{a}}\left(1 + \frac{1}{2}\Delta X_{gd}^{a} - \frac{1}{2}\Delta X_{gsd}^{s}\right)(V_{im} - k_c V_{tm}) - \frac{C_{gsd}^{s}}{C_{pg} + C_{gd}^{a}}\left(1 + \frac{1}{2}\Delta X_{gd}^{a} - \frac{1}{2}\Delta X_{gsd}^{s}\right)V_X + \frac{C_{gd}^{a}}{C_{pg} + C_{gd}^{a}}\left(1 + \frac{1}{2}\Delta X_{gd}^{a} - \frac{1}{2}\Delta X_{gd}^{a}\right)V_{DD} \tag{A4}$$

Taking the difference of these two equations yields

$$V_{od} = \frac{C_{gsd}^{s} + C_{pg}}{C_{pg} + C_{gd}^{a}}V_{id} + \frac{C_{gsd}^{s} + C_{pg}}{C_{pg} + C_{gd}^{a}}\left(-\Delta X_{gd}^{a} + \Delta X_{gsd}^{s}\right)V_{ic} - \frac{C_{gsd}^{s} + C_{pg}}{C_{pg} + C_{gd}^{a}}k_c\Delta V_T - \frac{C_{gsd}^{s} + C_{pg}}{C_{pg} + C_{gd}^{a}}k_c V_T\left(-\Delta X_{gd}^{a} + \Delta X_{gsd}^{s}\right) - \frac{C_{gsd}^{s}}{C_{pg} + C_{gd}^{a}}\left(-\Delta X_{gd}^{a} + \Delta X_{gsd}^{s}\right)V_X \tag{A5}$$

Rearranging the terms gives

$$V_{od} = \frac{C_{gsd}^{s} + C_{pg}}{C_{pg} + C_{gd}^{a}}\left(V_{id} - k_c\Delta V_T - \left(\Delta X_{gd}^{a} - \Delta X_{gsd}^{s}\right)(V_{ic} - k_c V_T)\right) - \frac{C_{gsd}^{s}}{C_{pg} + C_{gd}^{a}}\left(-\Delta X_{gd}^{a} + \Delta X_{gsd}^{s}\right)V_X$$

$$V_{od} = \frac{C_{gsd}^{s} + C_{pg}}{C_{pg} + C_{gd}^{a}}\left[V_{id} - k_c\Delta V_T - \left(\Delta X_{gd}^{a} - \Delta X_{gsd}^{s}\right)(V_{ic} - k_c V_T) - \frac{C_{gsd}^{s}}{C_{gsd}^{s} + C_{pg}}\left(-\Delta X_{gd}^{a} + \Delta X_{gsd}^{s}\right)V_X\right] \tag{A6}$$

Finally, separating V_t and capacitance imbalances from one another, we arrive at the following final equation

$$V_{od} = \frac{C_{gsd}^s + C_{pg}}{C_{pg} + C_{gd}^a}\left[V_{id} - k_c\Delta V_t - \left(\Delta X_{gd}^a - \Delta X_{gsd}^s\right)\left(V_{ic} - k_c V_{t\_avg} - \frac{C_{gsd}^s}{C_{gsd}^s + C_{pg}}V_X\right)\right] \tag{A7}$$

$$where\ k_c = \frac{C_{gsd}^a + C_{pg}}{C_{gsd}^s + C_{pg}}$$

Similarly, to find the common mode output, we simply take the average of V_{op} and V_{om} to give

$$V_{oc} = \frac{C_{gsd}^s + C_{pg}}{C_{pg} + C_{gd}^a}\left(V_{ic} - k_c V_{t\_avg}\right) + \frac{1}{4}\frac{C_{gsd}^s + C_{pg}}{C_{pg} + C_{gd}^a}\left(-X_{gd}^a + X_{gsd}^s\right)(V_{id} - k_c\Delta V_t) - \frac{C_{gsd}^s}{C_{pg} + C_{gd}^a}V_X + \frac{C_{gd}^a}{C_{pg} + C_{gd}^a}V_{DD}$$

$$V_{oc} = \frac{C_{gsd}^s + C_{pg}}{C_{pg} + C_{gd}^a}\left(V_{ic} - \frac{C_{gsd}^s}{C_{gsd}^s + C_{pg}}V_X - kV_{T\_avg} + \frac{C_{gd}^a}{C_{gsd}^s + C_{pg}}V_{DD} + \frac{1}{4}\left(X_{gsd}^s - X_{gd}^a\right)(V_{id} - k\Delta V_T)\right) \tag{A8}$$

Since node X is coupled with the input via the following capacitance ratio

$$V_X = \frac{C_{gsd}^s}{C_{gsd}^s + C_{px}}V_{ic} \tag{A9}$$

The V_{ic} term of (A8) can thus be simplified to give

$$\frac{C_{gsd}^s + C_{pg}}{C_{pg} + C_{gd}^a}\left(1 - \frac{C_{gsd}^s}{C_{gsd}^s + C_{pg}}\frac{C_{gsd}^s}{C_{gsd}^s + C_{px}}\right) = \frac{1}{C_{pg} + C_{gd}^a}\left(C_{pg} + \frac{C_{gsd}^s C_{px}}{C_{gsd}^s + C_{px}}\right) \tag{A10}$$

Substituting this result back in the V_{oc} equation, we arrive at our final expression for the common mode output

$$V_{oc} = \frac{C_{pg} + \frac{C_{gsd}^{s} C_{px}}{C_{gsd}^{s} + C_{px}}}{C_{pg} + C_{gd}^{a}} V_{ic} + G_{ideal}\left(-k_c V_{T\_avg} + \frac{1}{4}\left(X_{gsd}^{s} - X_{gd}^{a}\right)\left(V_{id} - k_c \Delta V_T\right)\right) + \frac{C_{gd}^{a}}{C_{pg} + C_{gd}^{a}} V_{DD} \tag{A11}$$

APPENDIX B – SETTLING ANALYSIS

The equation that governs the settling behavior of the dynamic amplifier during amplification is given by

$$\frac{dV_s}{dt} = \frac{I_d - I_b}{C_{Ltot}} \tag{B1}$$

The gate and source of the transistor are related via a capacitive ratio, thus

$$\frac{dV_g}{dt} = \frac{C_{gs}}{C_{gs} + C_{pg} + C_{gd}} \frac{dV_s}{dt} = k_{gs} \frac{dV_s}{dt} \tag{B2}$$

Consequently, the relationship between V_{gs} settling and output V_s settling is

$$\frac{dV_{gs}}{dt} = -\frac{C_{pg} + C_{gd}}{C_{gs} + C_{pg} + C_{gd}} \frac{dV_s}{dt} = -\frac{1}{G_{ideal}} \frac{dV_s}{dt} = \left(k_{gs} - 1\right) \frac{dV_s}{dt} \tag{B3}$$

The solution to this differential equation (DFE) is

$$V_{gs}(t) = V_{gs}(0) + \left(k_{gs} - 1\right)\left(V_s(t) - V_s(0)\right) \tag{B4}$$

Strong inversion solution

Expressing the drain current using squared-law equation and assuming that I_b is large enough that the transistor always stays in strong inversion {i.e. $I_b \approx K\left[V_{gs}(t_s) - V_t(t_s)\right]^2 = KV_{ov}^2(t_s)$ } we have

$$\frac{dV_s}{dt} = \frac{I_d - I_b}{C_{Ltot}} \approx \frac{K\left[V_{gs}(t) - V_t(t)\right]^2 - KV_{ov}^2(t_s)}{C_{Ltot}} \tag{B5}$$

Due to DIBL, V_t varies with time. Nevertheless, we approximate the rate of threshold voltage's change as proportional to rate of change of the output

$$\frac{dV_t}{dt} \approx k_D \frac{dV_s}{dt} = -k_D \frac{C_{Ltot}}{C_{pg} + C_{gd}} \frac{dV_{gs}}{dt} = -z\frac{dV_{gs}}{dt} \tag{B6}$$

Simulation shows this is a reasonable approximation. The following figure plots the factor z for a typical transient settling case of V_{gs} showing that z stays relatively constant during amplification.

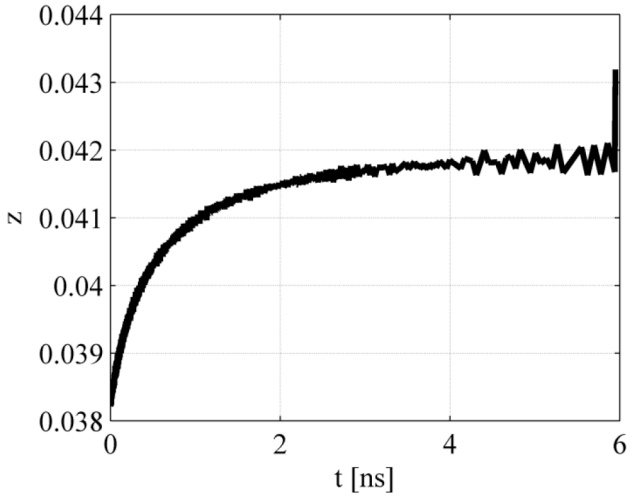


Fig. B.1 - Ratio of Vt's rate of change over Vout's rate of change

Thus, the solution of (B6) is given by

$$V_t(t) - V_t(0) = -z\left(V_{gs}(t) - V_{gs}(0)\right) = k_D\left(V_s(t) - V_s(0)\right) \tag{B7}$$

The overdrive voltage of the transistor can thus be deduced to give

$$V_{ov}(t) - V_{ov}(0) = \left(k_{gs} - k_D - 1\right)[V_s(t) - V_s(0)] = \frac{k_{gs} - k_D - 1}{k_{gs} - 1}\left[V_{gs}(t) - V_{gs}(0)\right] \tag{B8}$$

Re-expressing the I_d term of (B5), we have

$$K\left(V_{gs} - V_t\right)^2 = K\left\{\left(k_{gs} - 1\right)[V_s(t) - V_s(0)] + V_{gs}(0) - V_t\right\}^2 \tag{B9}$$

Substituting the expression for V_t from (B7) into (B9) leads to

$$\begin{aligned} K\left(V_{gs} - V_t\right)^2 = K\{&\left(k_{gs} - 1\right)[V_s(t) - V_s(0)] + V_{gs}(0) - V_t(0) \\ &- k_D\left(V_s(t) - V_s(0)\right)\}^2 \end{aligned} \tag{B10}$$

$$K\left(V_{gs} - V_t\right)^2 = K\left\{\left(k_{gs} - k_D - 1\right)[V_s(t) - V_s(0)] + V_{ov}(0)\right\}^2 \tag{B11}$$

Thus, rewriting (B5) using (B11), we have

$$\frac{C_{Ltot}}{K}\frac{dV_s}{dt} = \left\{\left(k_{gs} - k_D - 1\right)[V_s(t) - V_s(0)] + V_{ov}(0)\right\}^2 - V_{ov}^2(t_s) \tag{B12}$$

$$\frac{dV_s}{\left[V_s(t) - V_s(0) + \dfrac{V_{ov}(0)}{k_{gs} - k_D - 1}\right]^2 - \left[\dfrac{V_{ov}(t_s)}{k_{gs} - k_D - 1}\right]^2} = \frac{K}{C_{Ltot}}\left(k_{gs} - k_D - 1\right)^2 dt$$

$$\frac{dV_s}{(V_s + a)^2 - b^2} = \frac{dV_s}{{V_s}^2 + 2aV_s + a^2 - b^2} = \frac{K}{C_{Ltot}}\left(k_{gs} - k_D - 1\right)^2 dt$$

$$where \quad a = \frac{V_{ov}(0)}{k_{gs} - k_D - 1} - V_s(0) < 0, \qquad b = \frac{V_{ovf}}{1 - k_{gs} + k_D} > 0$$

Integrating both sides of the above equation yields

$$\frac{1}{2b} ln\left(\frac{V_s + a - b}{V_s + a + b} \cdot \frac{V_s(0) + a + b}{V_s(0) + a - b}\right) = \left(k_{gs} - k_D - 1\right)^2 \frac{K}{C_{Ltot}} t \tag{B13}$$

Note that

$$a \pm b + V_s(0) = \frac{V_{ov}(0) \mp V_{ov}(t_s)}{k_{gs} - k_D - 1}$$

Thus, the solution to the above DFE is

$$V_s(t) = \frac{1}{1 + k_D - k_{gs}} \left(\frac{2V_{ov}(t_s)}{1 - \frac{V_{ov}(0) + V_{ov}(t_s)}{V_{ov}(0) - V_{ov}(t_s)} e^{t/\tau}} + V_{ov}(0) - V_{ov}(t_s) \right) + V_s(0) \qquad \text{(B14)}$$

Or equivalently

$$V_s(t) = \frac{V_{ov}(t_s)}{1 + k_D - k_{gs}} \left(\frac{2}{1 - \frac{R_f + 1}{R_f - 1} e^{t/\tau}} + R_f - 1 \right) + V_s(0) \qquad \text{(B15)}$$

$$where \quad \tau = \frac{C_{Ltot}}{2K(1 + k_D - k_{gs})V_{ov}(t_s)} \quad and \;\; R_f = \frac{V_{ov}(0)}{V_{ov}(t_s)}$$

It then follows from (B4) that

$$V_{gs}(t) - V_{gs}(0) = \frac{k_{gs} - 1}{1 + k_D - k_{gs}} \left(\frac{2}{1 - \frac{R_f + 1}{R_f - 1} e^{t/\tau}} + R_f - 1 \right) V_{ov}(t_s) \qquad \text{(B16)}$$

Since $k_D << k_{gs}$, we can approximately write

$$V_{gs}(t) - V_{gs}(0) \approx -V_{ov}(t_s) \left(\frac{2}{1 - \frac{R_f + 1}{R_f - 1} e^{t/\tau}} + R_f - 1 \right) \qquad \text{(B17)}$$

Defining the V_{gs} settling error as

$$\varepsilon(t) = \frac{V_{gs}(t) - V_{gs}(\infty)}{V_{gs}(0) - V_{gs}(\infty)} \qquad \text{(B18)}$$

Here, we have

$$V_{gs}(\infty) - V_{gs}(0) = \frac{k_{gs} - 1}{1 + k_D - k_{gs}} \left(R_f - 1\right) V_{ov}(t_s)$$

$$V_{gs}(t) - V_{gs}(\infty) = \frac{k_{gs} - 1}{1 + k_D - k_{gs}} \frac{2}{1 - \frac{R_f + 1}{R_f - 1} e^{t/\tau}} V_{ov}(t_s)$$

Thus, using these 2 equations, the settling error is

$$\varepsilon(t) = \frac{V_{gs}(t) - V_{gs}(\infty)}{V_{gs}(0) - V_{gs}(\infty)} = \frac{2}{\left(R_f + 1\right) e^{t/\tau} - \left(R_f - 1\right)} \tag{B19}$$

If ε_{spec} is the desired V_{gs} error, the time it takes to reach this error is

$$t = \tau \cdot ln\left(\frac{R_f - 1}{R_f + 1} + \frac{2}{\varepsilon_{spec}} \frac{1}{R_f + 1}\right) \tag{B20}$$

Weak inversion solution: Using the same analysis step and re-expressing (B1) but with the exponential-law current for weak inversion operation, we have

$$\frac{dV_{gs}}{dt} = \left(k_{gs} - 1\right) \frac{dV_s}{dt} = \frac{k_{gs} - 1}{C_L} \left[I_o exp\left(\frac{V_{gs}}{nV_T}\right) - I_b\right] \tag{B21}$$

The solution to this equation is

$$V_{gs}(t) = -nV_T ln\left(\frac{I_o}{I_b} - R_o \exp\left(-t/\tau\right)\right) \tag{B22}$$

$$where \quad R_o = \frac{I_o}{I_b} - \exp\left(-\frac{V_{gs0}}{nV_T}\right) \quad and \quad \tau = G_{ideal} \frac{nV_T}{I_b} C_L$$

It thus follows from the above that the V_{gs} error is

$$\varepsilon(t) = \frac{V_{gs}(t) - V_{gs}(\infty)}{V_{gs}(0) - V_{gs}(\infty)} = \frac{ln\left(\frac{I_o}{I_b}\right) - ln\left(\frac{I_o}{I_b} - R_o \exp\left(-t/\tau\right)\right)}{\frac{V_{gs}(0)}{nV_T} + ln\left(\frac{I_o}{I_b}\right)}$$

$$\varepsilon(t) = \frac{V_{gs}(t) - V_{gs}(\infty)}{V_{gs}(0) - V_{gs}(\infty)} = -\frac{ln\left(1 - R_1 \exp\left(-{}^{t}/_{\tau}\right)\right)}{\frac{V_{gs}(0)}{nV_T} + ln\left(\frac{I_o}{I_b}\right)} \tag{B23}$$

$$where \quad R_1 = 1 - \frac{I_b}{I_o} \exp\left(-\frac{V_{gs0}}{nV_T}\right)$$

APPENDIX C – NOISE DUE TO INCOMPLETE SETTLING

The noise analysis in Section 3.5 neglects the effect of incomplete settling. In particular, by referring to Fig. C.1 which shows the equivalent noise circuit of the amplifier during amplification and neglecting Flicker noise, we have

$$C_s \frac{d\overline{v_n}}{dt} + \frac{\overline{v_n}}{R} = \overline{i_n} \tag{C.1}$$

$$where \quad \overline{i_n^2} = 4k_B T\gamma g_m \Delta f \ \ and \ R = (g_m + g_{mb} + r_o^{-1})^{-1}$$

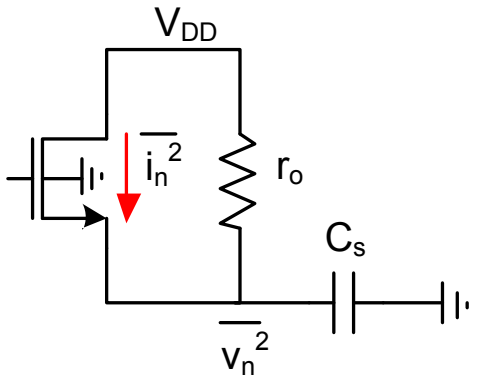


Fig. C.1 - Noise model during amplification

Solving the above equation is a complex task, since the noise current of the transistor is a function of g_m, which decays over time as the amplifier settles during amplification. Nevertheless, one can provide some bounds to the solution of this equation by respectively setting g_m to its minimum and maximum value (signifying the beginning and the end of amplification) and solving the equation with this constant g_m. This is a valid approach since the noise of the actual circuit is bounded between these two extremes. Thus, by assuming a constant g_m, equation (C.1) reduces to a form of Langevin's equation whose solution can be shown to be

$$v_n^2 = \frac{k_B T}{C_s}\left[e^{-2t_s/\tau} + \gamma g_m R\left(1 - e^{-2t_s/\tau}\right)\right]$$

$$\approx \frac{k_B T}{C_s}\left[e^{-2t_s/\tau} + \gamma \frac{g_m}{g_m + g_{mb} + r_o^{-1}}\left(1 - e^{-2t_s/\tau}\right)\right]$$

$$where \quad \tau = RC_s \ and \ t_s \ is \ the \ amplifier's \ settling \ time$$

This equation reduces to $k_B T/C_s$ for $\gamma = 1$ and large g_m.

APPENDIX D – SUMMARY OF ANALYTICAL ESTIMATES

The following table provides a summary of analytical estimates (through either simulation or computation) of the converter's performance

Table D. 1 - Summary of analytical estimates

Amplifier's gain $G = V_{od} / V_{id}$	1.55…1.65
Amplifier's settling time t_s	650…700 ps
Amplifier's linearity	> 50 dB
ADC quantization noise	0.74…1.56 mV
ADC input referred thermal noise	800…850 µV
ADC jitter induced noise (1…10 ps-rms)	0.1…1 mV
Total input referred noise	1.1…2.1 mV

BIBLIOGRAPHY

[1] B. Murmann, "EE315 Course Notes." Stanford University, 2010.

[2] B. Murmann, "Trends in Low-Power, Digitally Assisted A/D Conversion," *IEICE Trans. Electronics*, vol. E93-C, no. 6, pp. 718–729, Jun. 2010.

[3] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Select. Areas Comm.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[4] B. Murmann, "'ADC Performance Survey 1997-2011,' [Online]. Available: http://www.stanford.edu/~murmann/adcsurvey.html.".

[5] M. Horowitz, "EE313 Course Notes." Stanford University, 2010.

[6] N. H. E. Weste and D. M. Harris, *CMOS VLSI Design*, 4th ed. Adison-Wesley, 2011.

[7] G. Van der Plas and B. Verbruggen, "A 150MS/s 133µW 7b ADC in 90nm digital CMOS Using a Comparator-Based Asynchronous Binary-Search sub-ADC," in *ISSCC Dig. Tech. Papers*, San Francisco, CA, 2008, pp. 242–610.

[8] Hegong Wei, Chi-Hang Chan, U-Fat Chio, Sai-Weng Sin, U Seng-Pan, R. Martins, and F. Maloberti, "A 0.024mm2 8b 400MS/s SAR ADC with 2b/cycle and resistive DAC in 65nm CMOS," in *ISSCC Dig. Tech. Papers*, San Francisco, CA, 2011, pp. 188–190.

[9] Yung-Hui Chung and Jieh-Tsorng Wu, "A 16-mW 8-Bit 1-GS/s subranging ADC in 55nm CMOS," in *Symp. VLSI Circuits Dig.*, 2011, pp. 128–129.

[10] R. Nguyen, C. Raynaud, A. Cathelin, and B. Murmann, "A 6.7-ENOB, 500-MS/s, 5.1-mW dynamic pipeline ADC in 65-nm SOI CMOS," in *Proc. ESSCIRC*, Helsinki, Finland, 2011, pp. 359–362.

[11] J. Hu, N. Dolev, and B. Murmann, "A 9.4-bit, 50-MS/s, 1.44-mW pipelined ADC using dynamic residue amplification," in *Symp. VLSI Circuits Dig.*, 2008, pp. 216–217.

[12] E. Iroaga and B. Murmann, "A 12-Bit 75-MS/s Pipelined ADC Using Incomplete Settling," *IEEE J. Solid-State Circuits*, vol. 42, no. 4, pp. 748–756, Apr. 2007.

[13] M. Gustavsson, J. J. Wikner, and N. Tan, *CMOS Data Converters for Communications*, 2000th ed. Springer, 2011.

[14] F. Maloberti, *Data Converters*, 2007th ed. Springer, 2010.

[15] Y. Chiu, P. R. Gray, and B. Nikolic, "A 14-b 12-MS/s CMOS pipeline ADC with over 100-dB SFDR," *IEEE J. of Solid-State Circuits*, vol. 39, no. 12, pp. 2139– 2151, Dec. 2004.

[16] Y. C. Huang and T. C. Lee, "A 10b 100MS/s 4.5 mW pipelined ADC with a time sharing technique," in *ISSCC Dig. Tech. Papers*, San Francisco, CA, 2010, pp. 300–301.

[17] S. Devarajan, L. Singer, D. Kelly, S. Decker, A. Kamath, and P. Wilkins, "A 16b 125MS/s 385mW 78.7dB SNR CMOS pipeline ADC," in *ISSCC Dig. Tech. Papers*, 2009, pp. 86–87,87a.

[18] H.-S. Lee, B. M. Min, G. Manganaro, and J. W. Valvano, "A 14-b 100-MS/s Pipelined ADC With a Merged SHA and First MDAC," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 2613–2619, Dec. 2008.

[19] J. K. Fiorenza, T. Sepke, P. Holloway, C. G. Sodini, and H.-S. Lee, "Comparator-Based Switched-Capacitor Circuits for Scaled CMOS Technologies," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2658–2668, Dec. 2006.

[20] J. Chu, L. Brooks, and H.-S. Lee, "A zero-crossing based 12b 100MS/s pipelined ADC with decision boundary gap estimation calibration," in *IEEE Symp. VLSI Circuits Dig.*, 2010, pp. 237–238.

[21] B. Murmann and B. E. Boser, "A 12 b 75 MS/s pipelined ADC using open-loop residue amplification," in *ISSCC Dig. Tech. Papers*, San Francisco, CA, 2003, pp. 328– 497 vol.1.

[22] D.-L. Shen and T.-C. Lee, "A 6-bit 800-MS/s Pipelined A/D Converter With Open-Loop Amplifiers," *IEEE J. of Solid-State Circuits*, vol. 42, no. 2, pp. 258–268, Feb. 2007.

[23] A. Nazemi, C. Grace, L. Lewyn, B. Kobeissy, O. Agazzi, P. Voois, C. Abidin, G. Eaton, M. Kargar, C. Marquez, S. Ramprasad, F. Bollo, V. A. Posse, S. Wang, and G. Asmanis, "A 10.3GS/s 6bit (5.1 ENOB at Nyquist) time-interleaved/pipelined ADC using open-loop amplifiers and digital calibration in 90nm CMOS," in *Symp. VLSI Circuits Dig.*, 2008, pp. 18–19.

[24] J. K.-R. Kim and B. Murmann, "A 12-bit, 30-MS/s, 2.95-mW pipelined ADC using single-stage class-AB amplifiers and deterministic background calibration," in *Proc. ESSCIRC*, Sevilla, Spain, 2010, pp. 378–381.

[25] Hee-Cheol Choi, Young-Ju Kim, Myung-Hwan Lee, Young-Lok Kim, and Seung-Hoon Lee, "A 12b 50MS/s 10.2mA 0.18µm CMOS Nyquist ADC with a fully differential class-AB switched op-amp," in *Symp. VLSI Circuits Dig.*, 2008, pp. 220–221.

[26] M. Anthony, E. Kohler, J. Kurtze, L. Kushner, and G. Sollner, "A process-scalable low-power charge-domain 13-bit pipeline ADC," in *Symp. VLSI Circuits Dig.*, 2008, pp. 222–223.

[27] I. Ahmed, J. Mulder, and D. A. Johns, "A Low-Power Capacitive Charge Pump Based Pipelined ADC," *IEEE J. Solid-State Circuits*, vol. 45, no. 5, pp. 1016–1027, May 2010.

[28] P. R. Gray, P. J. Hurst, L. H. Stephen, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 5th ed. .

[29] A. N. Karanicolas, Hae-Seung Lee, and K. L. Barcrania, "A 15-b 1-Msample/s digitally self-calibrated pipeline ADC," *IEEE J. of Solid-State Circuits*, vol. 28, no. 12, pp. 1207–1215, Dec. 1993.

[30] J. Hu, "Low-power dynamic amplifiers for pipelined A/D conversion," Thesis, Stanford University, 2008.

[31] M. Miyahara, Y. Asada, Daehwa Paik, and A. Matsuzawa, "A low-noise self-calibrating dynamic comparator for high-speed ADCs," in *ASSCC*, 2008, pp. 269–272.

[32] M. A. Copeland and J. M. Rabaey, "Dynamic amplifier for MOS technology," *Electronics Letters*, vol. 15, no. 10, pp. 301–302, May 1979.

[33] K. Kotani, T. Shibata, and T. Ohmi, "CMOS charge-transfer preamplifier for offset-fluctuation cancellation in low-power A/D converters," *IEEE J. of Solid-State Circuits*, vol. 33, no. 5, pp. 762–769, May 1998.