Design Techniques to Improve Noise and Linearity of Data Converters

A Dissertation Presented
by

Haiyang Zhu

to

The Department of Electrical and Computer Engineering

in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy

in

Electrical Engineering

Northeastern University
Boston, Massachusetts

April 2016
To my family:

Jie, my lovely wife, for her love, support, and inspiration

Alice and Lily, my lovely daughters

To my beloved parents, Xingyin and Cuilan, who always support me
Contents

List of Figures iv
List of Tables vii
Acknowledgments viii
Abstract of the Dissertation ix

1 Introduction 1
1.1 Applications of Data Converters ................................................. 1
1.2 Performance of Data Converters ............................................... 3
1.2.1 System Performance .......................................................... 3
1.2.2 Digital Imaging Applications .............................................. 3
1.2.3 Wireless Communication Applications .................................. 5
1.3 Motivation and Scope ........................................................... 7
1.4 Outline ................................................................................. 8

2 Fundamentals of Data Converters 11
2.1 Introduction ............................................................................ 11
2.2 Quantization Error ................................................................... 11
2.3 Characterization of Data Converters ........................................... 14
2.3.1 Transfer Function Performance Metrics ................................. 14
2.3.2 Spectral Performance Metrics .............................................. 16
2.4 ADC Architectures .................................................................. 18
2.5 ADC Figures of Merit ............................................................... 22
2.6 ADC Circuit Building Blocks .................................................... 23
2.7 DAC Architectures .................................................................. 23
2.8 DAC Circuit Building Blocks .................................................... 26

3 Sampling Circuits That Break the $kT/C$ Thermal Noise Limit 28
3.1 Introduction ............................................................................. 28
3.2 Background .............................................................................. 29
3.2.1 $kT/C$ Noise ...................................................................... 29
3.2.2 Prior Thermal Noise Reduction Work ................................... 31
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Typical edge node of IoT.</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Digital imaging signal path.</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>Images converted by low-noise and high-noise ADCs.</td>
<td>4</td>
</tr>
<tr>
<td>1.4</td>
<td>Wireless receiver evolution. (a) dual conversion. (b) single conversion.</td>
<td>6</td>
</tr>
<tr>
<td>1.5</td>
<td>ADC non-linearity.</td>
<td>6</td>
</tr>
<tr>
<td>1.6</td>
<td>Wireless transmitter evolution. (a) traditional superheterodyne implemented</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>with baseband DACs. (b) complex IF modulator with IF DACs. (c) direct RF</td>
<td></td>
</tr>
<tr>
<td></td>
<td>synthesis with RF DAC.</td>
<td></td>
</tr>
<tr>
<td>1.7</td>
<td>DAC non-linearity.</td>
<td>8</td>
</tr>
<tr>
<td>1.8</td>
<td>Design trade-offs of data converters.</td>
<td>8</td>
</tr>
<tr>
<td>2.1</td>
<td>Ideal transfer function of data converters. (a) an ADC. (b) a DAC.</td>
<td>12</td>
</tr>
<tr>
<td>2.2</td>
<td>Quantization error of a 3-bit ideal ADC.</td>
<td>13</td>
</tr>
<tr>
<td>2.3</td>
<td>Distribution of quantization noise.</td>
<td>13</td>
</tr>
<tr>
<td>2.4</td>
<td>An example of INL and DNL for an ADC.</td>
<td>15</td>
</tr>
<tr>
<td>2.5</td>
<td>An example of INL and DNL for an DAC.</td>
<td>15</td>
</tr>
<tr>
<td>2.6</td>
<td>Histogram for a grounded-input ADC.</td>
<td>16</td>
</tr>
<tr>
<td>2.7</td>
<td>An example of the spectral performance metrics (2048-point FFT).</td>
<td>17</td>
</tr>
<tr>
<td>2.8</td>
<td>Published data at ISSCC and VLSI from 1997 to 2015 for ADC BW versus S/NDR.</td>
<td>19</td>
</tr>
<tr>
<td>2.9</td>
<td>Flash ADC.</td>
<td>20</td>
</tr>
<tr>
<td>2.10</td>
<td>Pipeline ADC.</td>
<td>20</td>
</tr>
<tr>
<td>2.11</td>
<td>SAR ADC.</td>
<td>21</td>
</tr>
<tr>
<td>2.12</td>
<td>∆Σ ADC.</td>
<td>21</td>
</tr>
<tr>
<td>2.13</td>
<td>Time-interleaving ADC.</td>
<td>22</td>
</tr>
<tr>
<td>2.14</td>
<td>S/H circuit. (a) schematic. (b) timing diagram.</td>
<td>24</td>
</tr>
<tr>
<td>2.15</td>
<td>Switched-capacitor amplifier.</td>
<td>25</td>
</tr>
<tr>
<td>2.16</td>
<td>Segmented current-steering DAC.</td>
<td>25</td>
</tr>
<tr>
<td>2.17</td>
<td>Current steering cell.</td>
<td>26</td>
</tr>
<tr>
<td>2.18</td>
<td>Two current sources consisting of identical PMOS transistors.</td>
<td>26</td>
</tr>
<tr>
<td>3.1</td>
<td>Basic sampling circuit. (a) Single transistor as sampling switch. (b)</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Equivalent noise circuit during track phase.</td>
<td></td>
</tr>
</tbody>
</table>
3.2 Noise power spectral density. (a) Impact of different sampling switch resistance. (b) Impact of aliasing due to sampling. 31
3.3 Correlated double sampling. (a) Functional diagram of CCD pixel including reset transistor. (b) Pixel output versus time, including noise from reset transistor. 32
3.4 Active noise cancellation of sampled noise. (a) Configured as a track-and-hold amplifier. (b) Switch control timing diagram. 33
3.5 Use of track-and-hold in signal chain with sub-ADC, including clock timing diagram. 34
3.6 Measured sample phase noise on test chip versus settling time, with varying size of auxiliary capacitor $C_2$. 37
4.1 Typical switched-capacitor track-and-hold amplifier. 40
4.2 Circuit model for signal analysis of the conventional THA during the amplification phase. 42
4.3 Circuit model for noise analysis of the conventional THA during the amplification phase. 42
4.4 Time allocation of sub-phases in the amplification phase for the switched-bandwidth amplifier. 43
4.5 Transient response of the switched-bandwidth THA and the conventional THA with $\omega_p = 6.93/T_s, \omega_{p1} = 1.5\omega_p, \omega_{p2} = 0.5\omega_p$ and $T_1 = T_2 = 0.5T_s$. 44
4.6 Ratio of output noise power of the switched-bandwidth THA to that of the conventional constant bandwidth THA vs. $\omega_{p2}/\omega_p$, for three combinations of $T_1$ and $T_2$. 46
4.7 Schematic of a conventional two-stage OTA. 47
4.8 Cascode NMOS in the transconductance stage of OTA. (a) schematic. (b) transient response of internal nodes. 47
4.9 The proposed OTA. (a) schematic. (b) timing diagram of the amplification phase. 48
4.10 Simulated waveforms of differential output ($V_{out}$) and drain of $M_{22}$ ($V_X2$), when the proposed THA settles to $-800$ mV. 49
4.11 Simulated $P_{n,a}$ of the proposed THA vs. time in amplification phase. (a) $P_{n,a,const}$ and $P_{n,a,switch}$. (b) ratio of $P_{n,a,switch}$ to $P_{n,a,const}$. 53
4.12 Die photograph. 54
4.13 Measured INLs of four signal channels operating at 90 MS/s. (a) Constant mode. (b) Switched mode. (c) Summary of peak-peak INL. 55
4.14 Measured histograms of one channel output, with mean value removed. 56
4.15 Measured average $P_{n, const}$ of four channels vs. bandwidth setting. 57
5.1 Two current sources consisting of identical transistors. 60
5.2 Simulated $g_m$ vs. temperature of a current source in 65nm CMOS process. 60
5.3 Monte Carlo simulation results of normalized current mismatch. 61
5.4 Calibrated Current Source (CCS) of conventional calibration technique [Mercer JSSC 2007]. 62
5.5 Calibration loop configuration of conventional calibration technique [Mercer JSSC 2007]. 63
5.6 Schematic of CCS in [Mercer JSSC 2007]. 64
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Calibrated Current Source (CCS) of $g_m$ calibration technique</td>
<td>66</td>
</tr>
<tr>
<td>6.2</td>
<td>Schematic of the new CAL DAC</td>
<td>67</td>
</tr>
<tr>
<td>6.3</td>
<td>Schematic of $g_m V_{ref}$ current generator</td>
<td>68</td>
</tr>
<tr>
<td>6.4</td>
<td>Simulation results of the temperature drift of 15 MSB CCSs. (a) no calibration. (b) $g_m$ calibration technique in [Mercer JSSC 2007]. (c) $g_m$ calibration technique</td>
<td>70</td>
</tr>
<tr>
<td>6.5</td>
<td>Simulation results of INL and DNL v.s. temperature. (a) no calibration. (b) calibration technique in [Mercer JSSC 2007]. (c) $g_m$ calibration technique</td>
<td>71</td>
</tr>
<tr>
<td>7.1</td>
<td>Calibrated Current Source (CCS) of the proposed two-parameter calibration</td>
<td>74</td>
</tr>
<tr>
<td>7.2</td>
<td>An example of the proposed two-parameter calibration procedure</td>
<td>75</td>
</tr>
<tr>
<td>7.3</td>
<td>Schematic of CAL DACs I and g.</td>
<td>77</td>
</tr>
<tr>
<td>7.4</td>
<td>Calibration loop configuration</td>
<td>78</td>
</tr>
<tr>
<td>7.5</td>
<td>Measured 15 MSB CCSs vs. temperature and the worst standard deviation across the temperature range</td>
<td>79</td>
</tr>
<tr>
<td>7.6</td>
<td>Measured convergence of CAL DAC input codes of 15 MSB CCSs. (a) CAL DAC I. (b) CAL DAC g.</td>
<td>82</td>
</tr>
</tbody>
</table>
List of Tables

4.1 Measured standard deviation of four channel outputs referred to THA output . . . . 56
4.2 Simulated ratio of $P_{n,a}$ of the reduced bandwidth setting to $P_{n,a,max}$ . . . . . 56
6.1 Summary of INL and DNL ($-40^\circ$C-120$^\circ$C) . . . . . . . . . . . . . . . . . . . . . . 71
7.1 Measured temperature drift of 15 MSB CCSs . . . . . . . . . . . . . . . . . . . . . . . 81
Acknowledgments

Here I wish to thank those who made it possible for me to complete this dissertation. First and foremost, I would like to express my sincere appreciation to my advisor Professor Yong-Bin Kim of Northeastern University. Prof. Kim provided me with the guidance, wisdom, and encouragement to pursue and successfully complete this body of work.

I would also like to thank my Ph.D. committee members, Prof. Nian X. Sun, and Dr. Gabriele Manganaro for their valuable comments and recommendations regarding this thesis.

This research work was supported by Analog Devices, Inc. I would like to thank David Robertson, the General Manager for the High Speed Converter Group, and Gabriele Manganaro, the Engineering Director for the High Speed Converter Group, for their strong support.

A special of thanks goes to my colleagues, Ron Kapusta, Gil Engel, Wenhua Yang, and Nathan Egan, for the cooperation of this research work. I would like to thank Steve Rose and Martin Clara for their help on reviewing my publications.

I thank my lovely wife Jie for her love and inspiration throughout this long endeavor. Last, but not least, I thank my parents for supporting me always.
Abstract of the Dissertation

Design Techniques to Improve Noise and Linearity of Data Converters

by

Haiyang Zhu

Doctor of Philosophy in Electrical Engineering

Northeastern University, April 2016

Dr. Yong-Bin Kim, Adviser

The data converters including analog-to-digital converters (ADCs) and digital-to-analog converters (DACs) act as interfaces between a DSP-based system and the physical analog world. They are used in a wide variety of applications such as sensing, control and communication. Sensors convert the physical signal such as pressure, temperature, gas, speed, acceleration, and light into analog electronic signals that are digitized by ADCs for digital processing afterwards. DACs are used in transforming the results of digital processing, back to physical word for control, information display, or further analog processing. In a wireless infrastructure, the mobile handset communicates via the base station (BTS) with another mobile device. ADCs and DACs are critical blocks in the receiver and transmitter signal. Over the past decades, there is strong demand to improve the performance of data converters.

The ADCs and DACs are composed of basic circuit building blocks that face the design trade-offs between noise, linearity and power consumption. Switched-capacitor circuits are the most popular approach for realizing discrete-time analog signal processing in CMOS processes. They are widely used in ADCs as the sample-and-hold circuits and the amplification blocks, determining the noise and linearity of the entire ADC. Higher power consumption is usually required to lower the noise. For a current-steering DAC, the mismatch of the current sources determines its linearity. Calibration techniques improve the current mismatch without increasing the device area, enhancing both static linearity and dynamic performance. However, the conventional foreground calibration techniques is incapable of tracking temperature variations. This thesis proposes several design techniques to break the trade-offs by providing new degrees of freedom in design and improve the noise and linearity without sacrificing the power consumption.

Two design techniques are proposed to improve the noise performance of the switched-capacitor amplifier. The first technique actively cancels the thermal noise sampled on an input
capacitor during the sample phase. The second technique reduces the noise through bandwidth-switching during the amplification phase. Measurements from the test chips demonstrate that the two design techniques reduce the noise power in each phase by 67% and 45%, respectively, when compared with the conventional amplifier.

Two foreground calibration techniques, $g_m$ calibration and two-parameter calibration, are developed to improve the current source matching and hence the linearity of a current-steering DAC. $g_m$ calibration introduces a calibration DAC (CAL DAC) that tracks the major temperature variations of the mismatch. Two-parameter calibration further improve the temperature stability by using two CAL DACs to compensate two current source mismatch components respectively. The sum current of the two CAL DACs automatically tracks the exact temperature variations of the mismatch. All the proposed design techniques were implemented in test chips and validated by measurement results. The matching of the currents is improved from intrinsic 12-bit to 16-bit across the temperature range from $-40^\circ C$ to $85^\circ C$, which shows superior temperature stability compared to the previously published schemes.
Chapter 1

Introduction

The digital signal processing (DSP) has revolutionarily changed the way that the information is processed. However, the physical signals are analog and data converters are essential gateways between the analog domain and the digital domain. The advantages of digital technology are only as good as the ability of the data converters that faithfully convert between the analog signal and the digital signal.

Over the past decades, there is strong demand to improve the performance of data converters. First, the increasing system performance, for example, the continuing expansion of broadband communication, requires high performance data converters. Second, DSP is a cost-effective method over its analog counterpart, and therefore it is desirable to push more signal processing into the digital domain. As a consequence, demanding requirements are put on the data converters. From another point of view, the performance improvements of data converters create new applications. Both “technology push” and “market pull” impel converter innovation [1].

1.1 Applications of Data Converters

The data converters including analog-to-digital converters (ADCs) and digital-to-analog converters (DACs) act as interfaces between a DSP-based system and the physical analog world. They are used in a wide variety of applications such as sensing, control and communication. Sensors convert the physical signals such as pressure, temperature, gas, speed, acceleration, and light into analog electronic signals that are digitized by ADCs for digital processing afterwards. DACs are used in transforming the results of digital processing back to physical word for control, information display, or further analog processing.
In a wireless infrastructure, a mobile handset communicates via the base station (BTS) with another mobile device. ADCs and DACs are critical blocks in the receiver and transmitter signal chains [2]. The wireless communication has been experiencing tremendous growth and will continue to do so in the future. The innovations of data converters allow performance and architecture advances of wireless communication.

The Internet-of-Things (IoT) is emerging as the next technology wave. The concept is to connect any device with a sensor or controller to the Internet (and/or to each other), including cell phones, appliances, cars, components of machines and wearable devices. The analyst firm Gartner predicts that by 2020 there will be over 25 billion connected devices [3]. IoT is comprised of three conceptual elements: the edge nodes, the gateway nodes, and the cloud. An edge node is the “thing” providing an interface between the digital world of the Internet and the real analog world. The functionality of an edge node can be described in Figure 1.1. The transducers transform real world information to electrical signals, and vice versa (e.g., temperature, pressure, blood chemistry, or brain waves). ADCs and DACs act as interfaces between analog transducer and digital micro-controller. Wireless transceivers send or receive information between the “thing” and the network through data converters. Therefore, data converters are critical components in the edge node.

Figure 1.1: Typical edge node of IoT.
1.2 Performance of Data Converters

1.2.1 System Performance

The capability of information processing in a system is determined by two fundamental dimensions: bandwidth and dynamic range, as stated by Shannon-Hartley theorem [4]

\[ C = BW \log_2 (1 + DR) \],

where \( C \) is the information in bits per second; \( BW \) is the bandwidth of the system; \( DR \) (dynamic range) is the ratio of signal to unwanted error signals including noise and spurious signals, which is a generalized SNR (signal-to-noise ratio). For data converters, spurious signals are caused by their non-linearity, which reflects how faithful when the signal is converted between the analog domain and the digital domain.

Shannon’s theorem applies in a variety of applications besides communication [5]. For example, in an imaging system, the bandwidth means how many pixels can be processed in a given amount of time, while the dynamic range indicates the intensity between the dimmest and brightest light source that the system can handle.

The system bandwidth is usually fixed in the applications. For example, the wireless communication standards assign the corresponding bandwidth. Therefore, the noise and linearity determining DR performance are of particular interest in this thesis. In this chapter, the impact of the noise and linearity on two popular applications are described.

Power consumption is another important specification, especially for portable and battery-powered applications. To prolong the battery lifetime while keeping the device at reasonable size, mobile devices require power optimization at all levels of the system hierarchy. Traditionally the wireless communication infrastructure focuses on DR performance, whereas the associated power consumption is a lower priority. However, lower power consumption is strongly desired in order to save the energy cost. In addition, the trend of reduced size of base station requires low power consumption to simplify heat dissipation design and reduce the cost.

1.2.2 Digital Imaging Applications

For digital imaging applications, a typical signal path is shown in Figure 1.2. Usually three components are involved including an image sensor, an ADC, and a digital image signal processor. The image sensor converts light to electrical analog signal. The ADC performs the analog-to-digital
CHAPTER 1. INTRODUCTION

Figure 1.2: Digital imaging signal path.

Figure 1.3: Images converted by low-noise and high-noise ADCs.

conversion. The digital image signal processor further processes the raw digital data and generates the digital image.

The ADC non-linearity creates image artifacts noticeable to human eyes. If the Differential Non-Linearity (DNL) is large or there is abrupt transition in Integral Non-Linearity (INL) equivalently, it could transform smooth changes to “steps”. But if the transfer function of the INL is smooth, the human eyes are less sensitive to the moderate errors.

The ADC noise directly affects the dynamic range of the imaging system. Dynamic range is determined by comparing the maximum signal that can be processed to the minimum signal level that can be resolved in the system. The ADC noise consists of wide-band noise from its analog signal processing circuitry plus the quantization noise determined by its resolution. Noise degrades the image quality and creates “snow” or random variation in image. Therefore, the noise generated within the ADC must be minimized particularly in high quality imaging applications such as digital single-lens reflex camera (DSLR). Figure 1.3 shows, side by side, two images converted by low-noise and high-noise ADCs. Note that the ADC noise has a big impact on the image quality.
1.2.3 Wireless Communication Applications

Along the wireless receivers and transmitter’s signal chains, the ADCs and DACs locate at the boundary between analog and digital processing. The communication standards such as GSM, WCDMA, and LTE determine the minimum SNR level to allow the signal to be demodulated. In order to handle multiple user channels and multiple standards, the converters have shifted steadily toward the antenna in both receiver and transmitter signal chains to provide higher flexibility, integration, and lower overall costs [6][7]. The data converters are demanded for larger bandwidth and higher dynamic range, so become the bottleneck of the entire system performance.

First consider this shift in the receiver side. Low-pass ADCs with bandwidths in the kHz range were initially used to digitize a single channel or carrier at baseband, as depicted in Figure 1.4. The interfering signals (blockers) could be attenuated easily in the radio signal chain by analog filtering with a surface acoustic wave (SAW) filter. When only the wanted signal is present at the ADC input, the linearity requirement is not high. Currently, the analog-to-digital conversion is done at intermediate frequency (IF) by bandpass or sub-sampling ADCs, while the channels filtering is done by digital processing. Therefore, the ADCs must digitize the entire signal band with sufficient dynamic range to enable multi-carrier operation. The large blockers falling in the digitized band cannot be removed and desensitize the small wanted signals due to the ADCs’ non-linearity, as shown in Figure 1.5. Therefore, besides the noise, the ADCs’ non-linearity degrades the system DR. The future trend is that the ADCs are able to digitize GHz radio frequency (RF) signals directly, which places even higher performance requirements on the ADCs’ noise and linearity.

The same shift happens at the transmitter side as well. In the heterodyne type architecture as shown in Figure 1.6a, a pair of in-phase (I) and quadrature (Q) DACs drives baseband filtered data into a quadrature modulator. The modulators output is upconverted to RF frequency by one or two stages of mixers. Then a power amplifier (PA) amplifies the signal to the antenna. In the complex intermediate frequency (CIF) architecture, as shown in Figure 1.6b, the digital modulated signals instead of the baseband filtered data are directly sent to the I and Q IF DACs, which simplifies the filtering requirements and thus enables lower-cost filters to be implemented. Currently, the advances of RF DACs, as shown in Figure 1.6c, make it possible to generate the desired signal entirely in the digital domain and synthesize directly at RF frequency. The signal is filtered to clean up the spectrum, and then is sent to the PA and the antenna. All signal processes are done in the digital domain, which eliminates the local oscillator (LO) leakage and the upconverter’s image. The modulator can be considered ideal. The board area is significantly reduced and the filtering requirements between the
CHAPTER 1. INTRODUCTION

Figure 1.4: Wireless receiver evolution. (a) dual conversion. (b) single conversion. (c) direct RF conversion.

Figure 1.5: ADC non-linearity.

DAC and modulator are completely eliminated.

One key performance parameter of the DAC is the spectral purity, i.e. the level of spurious
CHAPTER 1. INTRODUCTION

Figure 1.6: Wireless transmitter evolution. (a) traditional superheterodyne implemented with baseband DACs. (b) complex IF modulator with IF DACs. (c) direct RF synthesis with RF DAC.

frequency components present in the DAC output signal. Typically, spurs in the output spectrum are harmonically related to the input signal and are generated due to non-linearities in the DAC output response, as shown in Figure 1.7. The signals within one channel may give rise to harmonics in adjacent channels, causing corruption.

1.3 Motivation and Scope

A data converter is a system composed of many circuit building blocks. It is important to identify the building blocks dominating the noise and linearity performances of the whole data converter in order to improve them. However, most of the performances trade with each other, as illustrated in Figure 1.8.
CHAPTER 1. INTRODUCTION

Figure 1.7: DAC non-linearity.

Figure 1.8: Design trade-offs of data converters.

Many trade-offs arise from the fundamental physics laws. The thermal noise power added in the analog circuit is proportional to the sampling capacitance. Therefore, the capacitance has to increase by a factor of four in order to reduce the noise by half, which dictates four times power consumption to drive the capacitor. The device matching limits the linearity of data converters. Quadrupling the physical area typically reduces mismatch by 1 bit. The device area to yield very high resolution matching is impractical. Such trade-offs present many challenges in the design of data converters.

Therefore, it is worthwhile exploring the design space to break these tradeoffs. New design techniques on the circuit building blocks introduce new design freedoms. Certain performances can be improved without sacrificing the other performances such as power consumption.

1.4 Outline

The remainder of this thesis is organized as follows.
CHAPTER 1. INTRODUCTION

Chapter 2 presents the fundamentals of data conversion and also describes how the converters can be characterized. Typical architectures of the data converters are summarized. Critical circuit building blocks are described.

Chapter 3 describes a technique to reduce the noise in sampling phase of a switched-capacitor amplifier and break the so-called \( kT/C \) limit. The thermal noise sampled on an input capacitor is actively canceled using an amplifier, so that the noise at the amplifier output can be controlled independently of input capacitor size. Measurements from the test chip demonstrate sampled thermal noise power reduction of 67\%, respectively, when compared to conventional \( kT/C \)-limited sampling.

Chapter 4 proposes a technique to reduce the noise in amplification phase of a switched-capacitor amplifier. The proposed amplifier divides the amplification phase into two sub-phases. The first sub-phase starts with high bandwidth and high slew rate to approach the final value quickly. In the second sub-phase, the bandwidth can be significantly reduced, achieving the target settling accuracy but with much lower noise. This division allows the settling accuracy, noise, slew rate, and DC gain to be designed independently. Measurements from the test chip demonstrate that the proposed amplifier achieves 45\% noise power reduction in the amplification phase along with improved linearity and lower power consumption, when compared with the conventional amplifier.

Chapter 5 gives an analysis about the current source mismatch and its temperature dependence. The conventional calibration technique is introduced and its disadvantage of tracking temperature variations is discussed.

Chapter 6 develops a temperature insensitive calibration technique for the current-steering DACs. Each current source is in parallel with a calibration DAC (CAL DAC) injecting a small correction current that corrects the mismatch and tracks the temperature variations. High matching accuracy is not only achieved at the calibration temperature, but also maintained across a wide operating temperature range from \(-40^\circ C\) to \(120^\circ C\). A 14-bit DAC is designed in a standard 65 nm CMOS process to validate this calibration. The transistor level simulation results show that the new calibration technique reduces the worst case INL and DNL of the DAC across the whole temperature range by a factor of 15.7 and 12.8 compared with the intrinsic matching and by a factor of 2.9 and 2.8 compared with the conventional calibration method.

Chapter 7 presents a foreground DAC calibration technique that further improves temperature stability. Two CAL DACs provide correction currents to compensate two current source mismatch components caused by the threshold voltage mismatch and the current factor mismatch. Each CAL DAC has the same temperature dependence as its corresponding mismatch component.
CHAPTER 1. INTRODUCTION

The sum current of the two CAL DACs automatically tracks the temperature variations of the current source mismatch with no need of temperature information. To obtain the CAL DAC input codes, two different bias current settings are used instead of calibrating at two separate temperatures. This scheme significantly reduces the calibration time and cost. A 16-bit DAC with this calibration technique is fabricated in a 65 nm CMOS process. The measurement results show that the matching of the currents is improved from intrinsic 12-bit to 16-bit across the temperature range from $-40\,^\circ C$ to $85\,^\circ C$. The proposed calibration technique achieves superior temperature stability compared to the previously published schemes.

Chapter 8 gives concluding remarks regarding the design techniques proposed in this thesis and points out the future research directions.

This thesis interpolates material from five papers by the author [8, 9, 10, 11, 12]. Chapter 3 uses material from References [8] and [9]. Meanwhile, Chapter 4 is based on Reference [11]. Some material from References [10] and [12] has also been incorporated into Chapter 5. Chapter 6 is based on Reference [10]. Finally, Chapter 7 is based on Reference [12] and [13].
Chapter 2

Fundamentals of Data Converters

2.1 Introduction

Performance metrics are needed to quantify the noise and linearity performance of data converters. Traditionally, the resolution (number of bits) \( N \) and the sample frequency \( f_s \) are two specifications characterizing data converters, representing DR and BW respectively. However, other performance metrics are more meaningful depending on the specific applications [14].

The intrinsic quantization noise associated with the data converters is discussed in Section 2.2. The transfer function and spectral performance metrics are introduced in Section 2.3.

The data converter performances are closely linked with their implementations. Over the years, many architectures of data converters have been developed to achieve optimal performance specifications for different sampling rates and resolutions. Typical architectures are discussed with their associated performance tradeoffs in Section 2.4 and Section 2.7. The converters are composed of basic circuit building blocks. Section 2.6 and Section 2.8 identify the critical circuit building blocks that directly determine the converters’ performances such as noise, linearity, speed and power.

2.2 Quantization Error

An ideal ADC linearly maps an analog input \( V_{\text{in}} \) into a digital output \( D_{\text{out}} \), whereas an ideal DAC linearly maps a digital input \( D_{\text{in}} \) into an analog analog \( V_{\text{out}} \). Figure 2.1 shows the transfer function of an ideal ADC and an ideal DAC.

Quantization error is intrinsic to data converters. In the case of an ADC, the staircase-shaped transfer function causes the quantization error \( (e_q) \), illustrated in Figure 2.2 for a 3-bit ideal
ADC. $\Delta$ corresponds to quantization step size, also referred to as the least significant bit (LSB). The input information between the quantization step is lost \cite{15}. The quantization error grows out of
bounds beyond code boundaries. The full scale range ($V_{FS}$) is defined as the maximum input range that satisfies $|e_q| < \Delta/2$, which implies

$$V_{FS} = 2^N \cdot \Delta$$  \hspace{1cm} (2.1)

for a $N$-bit ADC.

Quantization error is a deterministic function of the signal. However it can be considered as “quantization noise” when the input signal spans a large number of quantization steps actively. The distribution of the quantization noise can be approximated as uniform white noise as shown in Figure 2.3.

The quantization noise power is given by

$$\overline{e_q^2} = \frac{1}{\Delta} \int_{-\Delta/2}^{\Delta/2} e_q^2 de_q = \frac{\Delta^2}{12}.$$  \hspace{1cm} (2.2)

If the input sinusoidal signal with an amplitude of $V_{FS}/2$ is fed to a $N$-bit ADC, the output
Signal-to-Quantization-Noise Ratio (SQNR) can be expressed as

$$SQNR = \frac{V_{FS}^2}{8 \Delta^2 / 12} = 1.5 \cdot 2^{2N} = 6.02N + 1.76(dB).$$  \hspace{1cm} (2.3)

An ideal DAC produces a perfect analog output based on the digital input and does not introduce quantization error between its input and output. The output full scale range of a $N$-bit DAC is well-defined as (2.1). However, the quantization noise (2.3) is still valid for a signal path with an ideal DAC because the DAC’s digital input is quantized in the digital domain.

### 2.3 Characterization of Data Converters

#### 2.3.1 Transfer Function Performance Metrics

The transfer function curve characterizes a data converter in a static sense. A real data converter can deviate from the ideal transfer function in many different ways. Typical errors are quantified by offset, gain error, DNL, and INL. The offset and gain errors are linear. They can be easily corrected or are not of interest for many applications. However, DNL and INL represent the non-linearity of the data converter and are very important for many applications as they are a source of harmonic distortions.

DNL and INL are defined as

- $DNL[k]$ of an ADC: The difference between the code bin width of code $k$ and the average code bin width.
- $DNL[k]$ of a DAC: The difference between the step size of code $k$ and the average step size.
- $INL[k]$ of an ADC: The deviation of the values on the actual transfer function from a straight line measured at transition level $k$.
- $INL[k]$ of a DAC: The deviation of the values on the actual transfer function from a straight line measured at step $k$.

The examples of INL and DNL for an ADC and a DAC are shown in Figure 2.4 and Figure 2.5.

Because the imaging signals are rarely pure sine waves, the noise of the ADC in digital imaging applications is often characterized in a static sense using a “grounded-input histogram” test [16]. The histogram of a number of output samples is measured with the input held at a dc voltage. The output is a distribution of codes, as shown in Figure 2.6, centered at the nominal value of the dc input. The standard deviation of the histogram corresponds to the noise rms value.
CHAPTER 2. FUNDAMENTALS OF DATA CONVERTERS

Figure 2.4: An example of INL and DNL for an ADC.

Figure 2.5: An example of INL and DNL for an DAC.
CHAPTER 2. *FUNDAMENTALS OF DATA CONVERTERS*

2.3.2 Spectral Performance Metrics

The transfer function performance metrics can not be directly applied to wireless communication applications. Communication systems are usually characterized in the frequency domain, hence spectral performance metrics should be used.

The most common metrics to represent the noise of an ADC are the signal-to-noise ratio (SNR)

\[
SNR = \frac{P_{\text{sig}}}{P_{\text{noise}}},
\]

(2.4)

where \( P_{\text{sig}} \) is the power of a single-tone sine-wave signal and \( P_{\text{noise}} \) is the power of the noise.

The noise spectral density (NSD) is a very important noise performance metric, which is a measurement of the noise per unit of bandwidth at a specified frequency.

Assuming that the NSD is flat over the signal BW, the relation between NSD and SNDR is

\[
NSD(\text{dB/Hz}) = -SNR(\text{dB}) - 10 \cdot \log(BW).
\]

(2.5)

Spurious free dynamic range (SFDR) indicates the linearity of a data converter, defined as the ratio of the signal to the worst spurious signal

\[
SFDR = \frac{P_{\text{sig}}}{P_{\text{ws}}},
\]

(2.6)

where \( P_{\text{ws}} \) is the power of the worst spur that may or may not be a harmonic of the original signal.

SFDR represents the smallest value of signal that can be distinguished from a large blocker in receiver and the spectral purity in transmitter, respectively. SFDR can be specified with respect to full-scale (dBFS) or with respect to the actual signal amplitude (dBC).
Another metric signal-to-noise-and-distortion ratio (SNDR) is defined as

\[ SNR = \frac{P_{\text{sig}}}{P_{\text{noise}} + P_h}, \]  

(2.7)

where \( P_h \) is the power of the harmonics. SNDR is especially important for a data converter in communication systems, since it measures the combined effect of noise, quantization error and harmonic distortions.

Effective number of bits (ENOB) is an equivalent metric to SNDR, as defined by

\[ ENOB = \frac{SNDR - 1.76}{6.02}. \]  

(2.8)

An 8-bit ADC is modeled with a third order distortion and an additive noise. The simulated spectral performance metrics with a full scale input sinusoid wave (0 dBFS) are shown in Figure 2.7.

In the case of DACs, the non-linearity and noise are usually characterized by SFDR and NSD, respectively.
2.4 ADC Architectures

A variety of data converter architectures cover different applications with tradeoffs of power, speed, and accuracy. The typical ADC architectures are:

- Flash ADCs.
- Folding ADCs.
- Pipeline ADCs.
- Successive Approximation Register (SAR) ADCs.
- ΔΣ ADCs.
- Time Interleaving (TI) ADCs.

Several architectures are briefly discussed in this section.

The ADC data published at the IEEE International Solid-State Circuits Conference (ISSCC) and the VLSI Circuit Symposium from 1997 until 2015 are collected by Prof. Boris Murmann at Stanford University \[17\]. Figure 2.8 shows a chart in the signal bandwidth $BW$ versus $SNDR$ space, where each architecture occupies a region of space indicating its optimum applications.

Flash ADCs perform A/D conversion by comparing the analog input with many reference values as shown in Figure 2.9. The results form $2^N$ thermometer-coded digital data representing $N$-bit accuracy. The advantages of flash ADCs are low latency and high speed. However, in a Flash ADC, the number of comparators is thus exponentially related to $N$, requiring larger area and more power to achieve increased resolution. Hence, flash ADCs are commonly used in low-resolution applications or as sub-ADCs of other ADC architectures.

A pipeline ADC is composed of several cascaded low-resolution stages to obtain high overall resolution as shown in Figure 2.10. Each stage performs coarse A/D conversion and passes the residue error to the next stage. The digital output codes of all stages are combined into $N$-bit digital output. Pipeline ADC architecture fits for low-power, high-speed and high-resolution applications. The drawback is high latency. Switched capacitor circuits are well suited for pipeline ADCs in CMOS processes. It is possible to combine the functions of sample-and-hold (S/H), subtraction, DAC, and gain into a single switched capacitor circuit, referred to as the Multiplying Digital-to-Analog Converter (MDAC).
CHAPTER 2. FUNDAMENTALS OF DATA CONVERTERS

SAR ADCs use a binary search algorithm, which is more component efficient than Flash ADCs. The topology of a typical SAR ADC is illustrated in Figure 2.11. The analog input is sampled by a sample-and-hold (S/H) circuit which operates at the Nyquist sampling rate. Then the input is compared with mid-scale. Based on the comparison result, the DAC adjusts its output. By successively repeating the same search for \( N \) times, the digital representation of the analog input can be determined to \( N \) bits. The significant advantage of the SAR ADC is that it does not require linear amplifiers but comparators, resulting in a compact area and simple design. In addition, the latency is only one clock cycle of the Nyquist-rate clock. However, the SAR ADC comes at the cost of restricted sample rate due to sequential search. SAR ADCs have traditionally been restricted to low to medium speed, and medium to high accuracy applications. However, SAR ADCs are recently moving to high-speed region due to fast speed and low power consumption in finer-lithography CMOS processes.

A \( \Delta \Sigma \) ADC includes a modulator and a digital decimation filter as shown in Figure 2.12.
The modulator measures the difference between the analog input signal and the analog output of a feedback DAC. A loop filter (integrator) then measures the difference and presents a sloping signal.
to the quantizer (coarse sub-ADC) that converts the loop filter’s output to a digital signal. The ∆Σ ADC topology uses oversampling to decrease the quantization noise in the band of interest. The decimation filter receives the input bit streams and gives one $N$-bit digital output depending on the over sampling ratio (OSR) value.

A block diagram of a Time-Interleaved (TI) ADC with $M$ sub-ADCs is shown in Figure 2.13. The analog input is sampled in a repetitive sequence at the Nyquist rate $f_s$. Then it is converted to an $N$-bit digital code by each sub-ADC at slower rate $f_s/M$. The $M$ $N$-bit outputs are combined.
to generate an $N$-bit digital output at sample rate $f_s$. A significant advantage of TI ADCs is that the sample rate $f_s$ increases linearly with the number of sub-ADCs ($M$). However, the performance of a TI ADC is limited by the mismatches between sub-ADCs. Therefore, TI ADCs work for the applications requiring extremely high sample rate and medium to low dynamic performance.

2.5 ADC Figures of Merit

The performance metrics summarized in Section 2.3 characterize physical parameters for a converter in specific applications. It is desired to measure power efficiency for a converter meeting the specifications and compare various converters which could differ widely in architecture. Figure of merit (FOM) is introduced for this purpose.

One of the most commonly used FOM is so called “Walden’s” FOM [18]:

$$FOM_W = \frac{P}{f_s \cdot 2^{ENOB}},$$  
(2.9)

where $P$ is the power dissipation of the ADC and $f_s$ is the sample rate. $FOM_W$ express the energy per conversion-step.
"Schreier’s” FOM [19] is a better representation of tradeoffs in thermal-noise limited designs. Its definition is

\[ FOM_{S} = SNDR(\text{dB}) + 10 \cdot \log\left(\frac{BW}{P}\right). \]  

(2.10)

2.6 ADC Circuit Building Blocks

A sample-and-hold (S/H) is used to sample an analog signal and to hold its value for some time [20], which converts a continuous-time input waveform to discrete-time values. The S/H block is located in the front-end and acts as a circuit building block for data converters such as SAR ADCs and pipeline ADCs, as described in Section 2.4. Track-and-hold (T/H) is the most common sampling scheme in high sample-rate ADCs. The two terms are used exchangeably in this thesis.

Switched-capacitor (SC) circuits are the most popular approach for realizing discrete-time analog signal processing in CMOS processes. The primary advantages of SC circuits include compatibility with CMOS process, good accuracy of time constants, good linearity and good temperature characteristics [21]. They can be used for a wide variety of functions such as gain and filtering.

Therefore, the S/H block is usually implemented using SC circuits in CMOS processes. A simple S/H can be implemented with a switch, a capacitor and a buffer, as shown in 2.14.

SC amplifiers can be used as S/H as well, and as blocks inside data converters for example the MDAC in a pipeline ADC. Hence, SC amplifiers determines the performances of the entire ADC such as noise and linearity. In the simplest case, the operation of a switch-capacitor amplifier takes place in two phases: sampling/tracking and amplification [22]. Additional noise is added to the signal during both phases and limits the SNR of most SC circuits. The additional noise power is in the form of \(kT/C\), where \(k\) is the Boltzmann constant, \(T\) is absolute temperature, and \(C\) is the sampling capacitance. The traditional methods increase the power consumption or the capacitor size (require more power to drive) to improve the noise performance. There exists a design tradeoff between power consumption and noise. Two design techniques are proposed in Chapter 3 and Chapter 4 to improve the noise performance without extra power consumption.

2.7 DAC Architectures

There exist various DAC architectures including capacitive DACs, resistive DACs, and current-steering DACs. The former two are usually used for precision applications or as internal
blocks of ADCs. The current-steering DACs are most popular for a wide range of applications especially high-speed applications. They are capable to directly drive low-impedance loads (resistors). In addition, they can be used as the feedback DAC in continuous-time ∆Σ ADCs, determining the ADCs’ linearity. Only current-steering DACs are considered in this thesis.

A current-steering DAC is composed of weighted current steering cells in parallel. Two segments are typically formed, i.e. a Most Significant Bits (MSB) segment and a Least Significant Bits (LSB) segment, as shown in Figure 2.16. M MSB bits are thermometer decoded, whereas the remaining L LSB bits are binary coded. A possible transistor-level implementation of MSB current-steering cell is shown in Figure 2.17, which scales for the LSB segment. In each current-steering cell, the current is provided by the cascoded current source. All cells share the same bias voltage \( V_{cs} \) and \( V_{cas} \). The current source feeds a pair of source-coupled switches. The digital inputs \( d \) and \( db \) from switch driver circuits (not shown in the figure) drive the switches to steer all current toward one of the two output nodes.
CHAPTER 2. FUNDAMENTALS OF DATA CONVERTERS

Figure 2.15: Switched-capacitor amplifier.

Figure 2.16: Segmented current-steering DAC.
2.8 DAC Circuit Building Blocks

The matching of the current sources determines the static linearity of a current-steering DAC. The current sources consisting of identically sized and biased PMOS transistors are shown in Figure 2.18.

The random device mismatch arising from the manufacture variance causes the mismatch of the current sources, which can be quantified using Pelgroms model [23].

\[
\sigma^2(\Delta V_{TH}) = \frac{A_{VT}^2}{W \cdot L} \tag{2.11}
\]

\[
\left(\frac{\sigma(\Delta \beta)}{\beta}\right)^2 = \frac{A_{\beta}^2}{W \cdot L} \tag{2.12}
\]
CHAPTER 2. FUNDAMENTALS OF DATA CONVERTERS

where $A_{VT}$ and $A_\beta$ are the Pelgrom constants and only dependent on the process; $W$ and $L$ are the width and length of the transitors.

A straightforward method to reduce the device mismatch is to increase the device physical area. Normally the standard deviation of the device mismatch decreases by a factor of two as the device area increases by a factor of four. However, the large device area leads to large die area and large parasitic capacitance, limiting the dynamic performance when a DAC operates in high frequency.

On the other hand, the calibration techniques improve the current mismatch without increasing the device area, enhancing both static and dynamic linearity simultaneously. A critical drawback of the conventional foreground calibration techniques is that they do not track varying operation conditions such as temperature, as mentioned in [24]. Two calibration techniques proposed in Chapter 6 and Chapter 7 exhibit improved temperature stability.
Chapter 3

Sampling Circuits That Break the $kT/C$

Thermal Noise Limit

3.1 Introduction

Switched capacitor circuits are the implementation of choice for many modern mixed-signal circuits, especially in CMOS technology. Inherent in any switched capacitor circuit are sampling operations; when a switch opens, freezing the charge on a capacitor, a sample is taken. In conjunction with amplifiers, the sampled charge can then be redistributed to other capacitors in order to implement a variety of circuit functions: buffers, gain blocks, filters, and data converters [25, 26]. At the circuit design level, one of the common issues with sampling is the addition of noise to a signal each time a sample is taken. This noise represents a major limitation on the performance of most switched-capacitor circuits.

While capacitors are noiseless circuit elements, the resistors or transistors used to transfer charge contribute noise. Typically, when considering the noise associated with sampling, thermal noise is the dominant noise source. Thermal noise, which is also called Johnson or Nyquist noise, occurs due to the random motion of carriers due to thermal agitation. Unlike many other noise sources, such as shot noise and flicker noise, thermal noise occurs in the absence of DC current flow. Therefore, even with a DC input and a sampling circuit that has reached thermal equilibrium, thermal noise will be present and limit the achievable signal-to-noise ratio (SNR). It should be noted that flicker noise can also be a significant noise contributor in switched capacitor circuits, particularly in the case of active circuits. However, noise that is slow-moving relative to the sample rate can be
reduced or eliminated using offset-cancellation techniques, such as an auto-zero configuration and amplifier chopping.

Focusing on the most basic example, a sample taken on a single capacitor \( C \) with a transistor acting as a switch, it can be shown that the total thermal noise power on the sampling capacitor is equal to \( kT/C \), where \( k \) is the Boltzmann constant, \( T \) is absolute temperature, and \( C \) is the sampling capacitance. While the details of the \( kT/C \) limit will be discussed in Section 3.2, there are significant implications of this limit. Specifically, in order to achieve lower noise in a sampled system, larger capacitors must be used. Unfortunately, when increasing capacitor size in order to lower noise, other performance parameters suffer. The impacts can include larger die area, higher power in the sampling stage, and higher power in the amplifier that drives these increased sampling capacitors. It would be desirable to possess an extra degree of freedom with which the sampled noise could be designed independently of sampling capacitor size. This chapter will discuss one technique that enables such a design degree of freedom.

The remainder of this chapter is organized as follows. Section 3.2 includes background information. Section 3.3 describes an active noise cancellation technique. Measurement results demonstrating each technique are included in the corresponding sections.

### 3.2 Background

#### 3.2.1 \( kT/C \) Noise

The total thermal noise power on a capacitor in parallel with a single resistor was first shown in to reduce to the familiar \( kT/C \) limit, using the equipartition theorem of thermodynamics. The analysis can be extended to a very simple track-and-hold sampling circuit, shown in Figure 3.1. Assuming that the sampling transistor is operating in the triode region with a small voltage potential between the drain and source, it can be represented by an equivalent noise generator whose noise power spectral density is equal to \( 4kTR_{ON}[V^2/Hz] \), where \( k \) is again the Boltzmann constant, \( T \) is absolute temperature, and \( R_{ON} \) is the on-resistance of the sampling switch. While this is equivalent to the circuit analyzed in, an alternative analysis uses Parseval’s Theorem to calculate the total thermal noise power on the sampling capacitor

\[
\bar{v_n}^2 = 4kTR_{ON} \times \int_0^\infty \frac{1}{1 + j2\pi fR_{ON}C} \left| \frac{1}{1 + j2\pi fR_{ON}C} \right|^2 df. \tag{3.1}
\]

For a simple single-pole system, such as Figure 3.1, it is often easiest to reduce the integral into an equivalent noise bandwidth (ENBW), and to then express the total noise power as the product
Figure 3.1: Basic sampling circuit. (a) Single transistor as sampling switch. (b) Equivalent noise circuit during track phase.

of the noise power spectral density and the ENBW as follows:

\[ ENBW = \int_0^\infty \left| \frac{1}{1 + j2\pi f R_{ON}C} \right|^2 df, \] (3.2)

\[ v_n^2 = 4kT R_{ON} \times \frac{1}{4R_{ON}C} = \frac{kT}{C}. \] (3.3)

In this simple case, the value of the transistor on-resistance \( R_{ON} \) appears in both the numerator (noise power spectral density) and the denominator (equivalent noise bandwidth). Therefore, the on-resistance terms cancel, and only the sampling capacitor remains as a degree of freedom in the expression for total noise power. This same “canceling” relationship is found to hold for more complicated structures as well, such as amplifiers in feedback that are used to provide a virtual ground for sampling.

The cancellation can be easily seen when the noise power spectral density sampled on the capacitor is plotted versus frequency, as in Figure 3.2. A lower transistor on-resistance decreases the thermal noise density, but the noise bandwidth is increased by the same ratio. There is no obvious way to decouple the inverse proportional relationship between the noise power density and the noise bandwidth.

The distinction between sampling bandwidth (or ENBW) and sample rate should be made clear and is also shown in Figure 3.2. Here, the track-mode ENBW, \( f_{\text{track}} \), is shown to be significantly larger than the sampling frequency, \( f_s \). In order to achieve good linearity, track-mode or settling bandwidth is often designed to be at least \( 0.25(N + 1) \) higher than the sampling frequency, where \( N \) is the number of bits of accuracy, and may be even higher in a sub-sampled system. However, the total sampled thermal noise is determined only by the noise spectral density and the equivalent noise bandwidth, and is not affected by the sample rate. Therefore, the remainder of this chapter will refer only to equivalent noise bandwidth and not sampling frequency.
CHAPTER 3. SAMPLING CIRCUITS THAT BREAK THE $kT/C$ THERMAL NOISE LIMIT

3.2.2 Prior Thermal Noise Reduction Work

Historically, image sensors such as CCDs have been very sensitive to $kT/C$ noise. One of the first techniques developed to mitigate the effect of thermal noise in image sensors was called correlated double sampling [33]. This technique is effectively a cancellation of thermal noise.
CHAPTER 3. SAMPLING CIRCUITS THAT BREAK THE KT/C THERMAL NOISE LIMIT

Figure 3.3: Correlated double sampling. (a) Functional diagram of CCD pixel including reset transistor. (b) Pixel output versus time, including noise from reset transistor.

associated with a reset transistor, and a functional diagram is shown in Figure 3.3. The key concept of this technique is that the thermal noise from the reset transistor, once sampled on capacitor $C_S$, will remain constant. Therefore, two readings of the pixel output can be taken, one just after the reset switch opens, and a second after the photocurrent has been integrated. Because the reset noise, $v_{ns}$, is correlated between these two readings, the final differenced output is independent of the sampled thermal noise from the reset switch. While the reset of a CCD pixel is a rather limited application, the concept of removing correlated noise is powerful.

More recently, several techniques have been proposed for CMOS image sensors that actively reset transistor noise [34, 35, 36]. The details of these techniques differ, but they operate on similar principles. A negative feedback loop is wrapped around the reset transistor, including an amplifier that controls one of its terminals. At frequencies for which the feedback loop has gain, the noise of the reset transistor is reduced by the negative feedback, typically requiring the bandwidth of the pixel reset to be limited. This is a compromise that can be tolerated in image sensors with relatively long reset times. Often, the bandwidth of the amplifier itself must also be restricted by using some auxiliary large capacitor; however, the amplifier and auxiliary capacitor do not reside inside the individual pixels, so the power and area penalty incurred are acceptable.

In contrast to these techniques that have been developed to counteract reset noise in image sensors, the circuits proposed in the following sections reduce or cancel sampled thermal noise while also being able to sample an arbitrary and time-varying input voltage. Also, because the proposed
CHAPTER 3. SAMPLING CIRCUITS THAT BREAK THE KT/C THERMAL NOISE LIMIT

Figure 3.4: Active noise cancellation of sampled noise (a) Configured as a track-and-hold amplifier. (b) Switch control timing diagram.

circuits include amplifiers, they are able to implement active switched capacitor functions, such as voltage gain or filtering.

3.3 Active Noise Cancellation

3.3.1 Circuit Configuration

The technique proposed here will use active circuits to cancel the thermal noise after it has already been sampled. An implementation of this technique is shown in Figure 3.4. The circuit blocks shown comprise a track-and-hold built using a two-stage amplifier with capacitive level-shifting between the two stages. The first amplifier $A_1$ is typically a low-gain pre-amplifier stage, though the technique would also work with a higher gain first stage.

Noise cancellation is achieved through appropriate design of the switch controls shown in Figure 3.4b. During the input sampling phase, both signals $\phi_1$ and $\phi_2$ are active (high). In this phase, the input voltage $V_{IN}$ is stored on input capacitor $C_S$, the offset of amplifier $A_1$ is stored on auxiliary capacitor $C_2$, and feedback capacitor $C_{FB}$ is cleared. When $\phi_1$ falls, the charge on the input and feedback capacitors is frozen. Thermal noise charge is also sampled with noise power.
equal to $kT / (C_S + C_{FB})$. Operation in this phase is identical to an output-referred auto-zero \cite{27}. This is also sometimes referred to as correlated double sampling \cite{37}, though the terms used as an offset cancellation technique can be confused with the application to reset noise cancellation \cite{33}, and is hence referred to herein as auto-zero.

After $\phi_1$ falls, thermal noise on $C_S$ is sampled and the summing node will settle to a final, noisy voltage. This voltage is amplified through $A_1$ and stored on capacitor $C_2$, as signal $\phi_2$ is still active. When $\phi_2$ falls, the sampled voltage on capacitor $C_2$ captures both the offset of amplifier $A_1$ and an amplified version of the thermal noise that was sampled at the summing node. Effectively, both offset in $A_1$ and the sampled thermal noise at the summing node will be auto-zeroed out via the same mechanism during phase $\phi_3$.

During $\phi_3$, the circuit is configured in hold mode and the input charge is transferred from $C_S$ to $C_{FB}$. However, the noise charge sampled on $C_S$ is not transferred to $C_{FB}$, and therefore does not appear at the amplifier output. To demonstrate that the noise charge is not transferred, it is easiest to begin with the assumption that $A_2$ has infinite gain. Therefore, the feedback loop will settle with no signal at the input to $A_2$. If the right-hand side of $C_2$ has no signal, then the left-hand side of $C_2$ and the summing node must remain at the same potential as they were when $\phi_2$ sampled. Therefore, the sampled thermal noise charge stays on the summing node and is not transferred to $C_{FB}$. A similar analysis can be used to show that the offset of $A_1$ also does not appear at the amplifier output.

It is also important to consider thermal noise sampled on capacitor $C_2$, as this noise is not cancelled during $\phi_3$. The sampled thermal noise power on $C_2$ is proportional to $kT/C_2$. One
obvious way to decrease this noise contributor is to increase capacitor \( C_2 \). While this increase may seem not desirable, \( C_2 \) is not driven by the input, and decreasing \( C_S \) at the expense of increased \( C_2 \) is often a favorable trade-off. Another approach to decreasing the noise contribution from the \( C_2 \) sample is to increase the gain of \( A_1 \), which lessens the impact of this noise when referred back to the input. The challenges with these approaches will be discussed next.

### 3.3.2 Impact on Sampling Bandwidth

A practical limitation that must be considered is the time required for \( A_1 \) to accurately amplify the noise sampled at the summing node. The time constant associated with the settling at the output of \( A_1 \) is determined by its output resistance, \( R_{OUT1} \), and capacitor \( C_2 \). For complete noise cancellation, the time allowed for settling, \( T_{DEL} \) in Figure 3.4(b), must be much longer than the settling time constant. The residual noise due to incomplete settling can be modeled as the exponential decay of a single pole system

\[
\frac{v_{n,\text{sample}}^2}{C_S} = kT \left( \frac{C_S + C_{FB} + C_P}{C_S} \right)^2 \left( e^{-\frac{2T_{DEL}}{R_{OUT1}C_S}} \right).
\]  

(3.4)

During the time between the falling edges of \( \phi_1 \) and \( \phi_2 \), the input voltage is still connected to the input of \( A_1 \) through the sampling capacitor. Any change in the input voltage is amplified at the output of \( A_1 \). Because this amplified version of the input is sampled on \( C_2 \), it is important that the amplified signal be linear in order to avoid distortion in the \( \phi_3 \) output. The requirement for linearity demonstrates why, while it is best for noise, a very large \( A_1 \) is not necessarily an optimal design trade-off.

In order to quantify the impact of \( A_1 \), the required input bandwidth must be defined. For an input sinusoid of amplitude \( V_{ampl} \) at a maximum input frequency \( f_{max} \), the maximum voltage change at the output of amplifier \( A_1 \) is

\[
\Delta v_{max} = A_1 V_{ampl} 2\pi f_{max} T_{DEL}.
\]  

(3.5)

This voltage change \( \Delta v_{max} \) must be within the linear signal swing of the amplifier. Therefore, as \( T_{DEL} \) is increased in order to reduce noise from the \( C_S \) sample, or as the gain of \( A_1 \) is increased in order to reduce noise from the \( C_2 \) sample, the output swing requirements of \( A_1 \) are increased. The output swing requirement is also directly related to the input signal bandwidth. Finally, \( A_1 \) must be capable of providing transient current to \( C_2 \), a requirement that scales with input signal frequency and the value of capacitor \( C_2 \).
CHAPTER 3. SAMPLING CIRCUITS THAT BREAK THE $kT/C$ THERMAL NOISE LIMIT

For relatively slow moving input signals, the proposed noise cancellation technique can be particularly powerful. By using a long $T_{DEL}$ and large $C_2$, the total sample phase noise can be extremely small, regardless of the sampling capacitor size. This can be advantageous in the case of oversampled systems; a small input capacitance is easy to drive during a short track time, while the comparatively slow moving input allows for significant noise reduction without placing difficult requirements on $A_1$.

3.3.3 Prototype Implementation

In order to verify the active noise cancellation technique, the circuit shown in Figure 3.4a has been implemented and embedded in a signal chain similar to Figure 3.5. This test chip is fabricated in a 65 nm CMOS process. The ADC clock rate is 20 MSPS. The size of the sampling capacitor is 2.3 pF. Other than the sampling capacitor, capacitance at the summing node totals 2.4 pF. The time constant at the output of amplifier $A_1$ is 550 ps. Based on calculation, the $kT/C_S$-limited sampled thermal noise should be $84 \mu$V-rms. Note that this calculation only includes the noise at the summing node and represents a lower limit to the achievable total sample phase noise; if offset cancellation were required and the impact of auto-zero capacitor $C_2$ were included in the calculation, the sample phase noise would be higher.

In consideration of the power dissipation of this test chip, there is effectively no impact due to the use of noise cancellation. The first stage amplifier design is constrained primarily by the need to maintain stable closed loop operation during phase $\phi_3$. Given this amplifier design, the noise cancellation technique works well for $T_{DEL} \sim 1$ ns. As predicted by (3.5), the amplifier output swing requirement is reasonable for inputs at the Nyquist frequency of 10 MHz. However, to support significantly higher input frequencies, the first stage amplifier power dissipation would be increased.

Figure 3.6 shows data measured from test chip as time $T_{DEL}$ is swept. The two solid lines shown correspond to configurations varying the size of auxiliary capacitor $C_2$. With $T_{DEL} = 0.1$ ns, there is very little cancellation of noise, as amplifier $A_1$ does not have adequate time to respond to the thermal noise sampled at the summing node. The maximum noise, 100 $\mu$V-rms, is larger than $kT/C_S$ due to the additional noise contributed by the $C_2$ sample. As expected, as $T_{DEL}$ is increased, the measured noise decreases. Most of the noise cancellation benefit is achieved for $T_{DEL}$ equal to roughly two time constants, a result that can be predicted by (3.4).

Figure 3.6 also compares the measured results to the prediction of a simple model, shown by dotted lines. The model is based on (3.4) as well as calculations to refer noise sampled on $C_2$.
to the input. For $T_{DEL}$ less than 0.55 ns, the model prediction matches the measured results to within 5%. For longer $T_{DEL}$, there is a more significant discrepancy between the measured data and the model prediction. One possible explanation is that the noise sampled on $C_S$ is not completely cancelled. An alternative explanation is the presence of a noise source that is present in the prototype measurement but not accounted for in the model. The magnitude of this un-modeled noise source would be roughly 32 $\mu$V-rms, referred to the summing node. Unfortunately, test modes in silicon were not adequate to diagnose the root cause.

With regards to the overall sample phase noise power, test chip demonstrates a reduction of 67%, or 4.8 dB, from 96 $\mu$V-rms to 55 $\mu$V-rms. Considering only the thermal noise power sampled at the summing node, it is effectively cancelled by at least 85%, from 84 $\mu$V-rms to 32 $\mu$V-rms. It is relevant to note the significant cancellation of noise sampled at the summing node, as the additional noise from the second sample at the output of $A_1$ may be reduced much further for applications in which the input signal bandwidth is limited.
3.4 Conclusion

While it has been commonly accepted as a fundamental limit of thermal noise when sampling on a capacitor, $kT/C$ is, in fact, not a limit at all. This chapter presented the circuit-level sampling technique that allows the size of the input capacitor to be determined almost independently of the noise requirement. The technique used active circuits and a second capacitor not driven by the input to cancel the noise sampled on the input capacitor. Test chip measurements were presented to demonstrate that the effective sampled thermal noise can be reduced by as much as 67% without change to the input capacitor. This technique provides a powerful new degree of freedom in design, making possible circuits that are both low noise and easy to drive.
Chapter 4

Noise Reduction Technique Through Bandwidth Switching

4.1 Introduction

Switched-capacitor (SC) circuits are widely used in many signal-processing circuits such as amplifiers, filters, and data converters, especially in CMOS technology. The charge redistribution track-and-hold amplifier (THA) is a commonly used SC circuit, and a typical implementation is shown in Figure 4.1. The THA is controlled by two non-overlapping clock phases, $\phi_1$ and $\phi_3$. The falling edge of $\phi_2$ is slightly before that of $\phi_1$, which is the well-known bottom plate sampling technique \[38\]. During the tracking phase ($\phi_1$), the input voltage is acquired on the sampling capacitor $C_s$, and the actual sample is taken at the falling edge of $\phi_2$. During the amplification phase ($\phi_3$), the operational transconductance amplifier (OTA) forces the summing node $V_{sum}$ to be the virtual ground via negative feedback. The charge sampled on $C_s$ is transferred to $C_f$ and then sampled by the following SC circuit.

Additional noise is added to the signal during both phases and limits the signal-to-noise ratio (SNR) of most SC circuits. This chapter focuses on the reduction of thermal noise which is typically the dominant noise source, as flicker noise can be reduced by using large input devices, an auto-zero configuration, or amplifier chopping \[39\].

A thorough analysis of the thermal noise in the THA is provided in \[32\] and \[40\]. During the tracking phase, the thermal noise power on the sampling capacitor is $kT/C$, traditionally considered as the limit of the sampled noise. Recently several techniques \[8\] \[9\] have been developed...
CHAPTER 4. NOISE REDUCTION TECHNIQUE THROUGH BANDWIDTH SWITCHING

![Typical switched-capacitor track-and-hold amplifier](image)

Figure 4.1: Typical switched-capacitor track-and-hold amplifier.

to break this limit. However, these techniques do not apply to the noise added in the amplification phase, which is usually dominated by the OTA.

A number of alternative OTA circuits have been reported to reduce power consumption. Open-loop amplifiers, as well as their use with incomplete settling, and integrator-based amplifiers were proposed in [41] [42] [43]. Comparator-based switched-capacitor (CBSC), zero-crossing based (ZCB) circuits and ring amplifiers were developed in [44] [45] [46] [47]. While these circuits are more power efficient than the feedback OTA, they lose some of its benefits such as insensitivity of closed loop gain to variations of circuit parameters. Some techniques require sophisticated digital calibration, complicating the circuit design. Dynamic amplifiers [48] [49] [50] are another design option, in which the bias current is reduced as the output settles. However, the bias current in [49] is almost zero when the amplifier settles, which worsens the noise performance.

The proposed THA in this chapter is also based on the closed-loop feedback OTA, but it decouples the performance tradeoffs by exploiting their time-dependent nature. It is observed that the output signal and noise are important only at the sampling instant of the following SC circuit, i.e. at the end of the amplification phase. Low small-signal bandwidth is desirable for low noise at the sampling instant, but that presents a challenge for achieving the desired settling accuracy. The proposed THA divides the amplification phase into two consecutive sub-phases with different bandwidths [51]. During the first sub-phase, it operates with high bandwidth to approach the final value quickly. Then, it significantly reduces the bandwidth in the second sub-phase, achieving the required settling accuracy but with much lower noise. This technique allows the settling accuracy and
CHAPTER 4. NOISE REDUCTION TECHNIQUE THROUGH BANDWIDTH SWITCHING

noise to be designed independently. Furthermore, the proposed implementation is a modification of a simple cascode structure with improved large-signal slew rate to accelerate settling. The measurement results show that the proposed THA reduces the noise power in the amplification phase by 45%, in addition to improvement in both power consumption and linearity. From another perspective, if the same noise target were achieved, the amplifier power consumption would be significantly reduced using the proposed technique.

The remainder of this chapter is organized as follows. Section 4.2 describes the noise reduction by switching the THA bandwidth. Section 4.3 describes a technique to improve the slew rate of the cascode output structure. Section 4.4 presents a circuit that combines the two techniques plus simulation results, followed by silicon measured results in Section 4.5.

4.2 Noise Reduction via Bandwidth Switching

4.2.1 Conventional THA

The circuit schematic of the conventional THA during the amplification phase $\phi_3$ is shown in Figure 4.2. It is modeled as a single-pole system with the pole at the output node. The resistance of the switches is neglected. The DC gain of the OTA is much larger than what is needed for the target accuracy. The feedback factor $\beta$ is $C_f/(C_f + C_s + C_{par})$, where $C_{par}$ is the parasitic capacitance at the summing node. The total transconductance of the OTA is $G_m$. The closed-loop signal transfer function (STF) is given by

$$ \frac{V_{out}}{V_i} = -c \left( 1 - \frac{s}{\omega_z} \right) \left( 1 + \frac{s}{\omega_p} \right) $$

(4.1)

where $c = C_s/C_f$ is the closed-loop DC gain, $\omega_z = G_m/C_f$ is the zero frequency, and

$$ \omega_p = \frac{\beta G_m}{C_L + (1 - \beta)C_f} $$

(4.2)

is the pole frequency. The 3-dB bandwidth ($f_{3dB}$) of the STF is defined as $\omega_p/(2\pi)$.

$V_{si}$ is the voltage sampled on $C_s$ at the end of the tracking phase. The node $V_i$ connects to ground at the beginning of $\phi_3$, representing a step input $-V_{si}$. The output $V_{out}$ settles to $V_s$ after a fixed settling time $T_s$ during $\phi_3$, where $V_s = V_{si} \cdot c$. The transient response can be obtained by Laplace transform as follows:

$$ V_{out}(t) = V_s \cdot \left( 1 - \left( 1 + \frac{\omega_p}{\omega_z} \right) e^{-\omega_p t} \right). $$

(4.3)
CHAPTER 4. NOISE REDUCTION TECHNIQUE THROUGH BANDWIDTH SWITCHING

Assuming $\omega_p/\omega_z \ll 1$, $V_{out}(t)$ is simplified as

$$V_{out}(t) \approx V_s \cdot (1 - e^{-\omega_p t}). \quad (4.4)$$

The settling error after $T_s$ is $V_s e^{-\omega_p T_s}$, which limits the THA accuracy. In order to achieve $N$-bit settling accuracy, the 3-dB bandwidth $f_{3dB}$ is required to meet

$$f_{3dB} \geq 0.693 N/(2\pi T_s). \quad (4.5)$$

The circuit model for noise analysis is shown in Figure 4.3. Only the OTA noise is considered while the switch noise is ignored. The noise contributions of all devices inside the OTA are lumped together and referred to the OTA input. The power spectral density (PSD) of the input-referred noise is defined as $S_{eq}(f)$. Typically the noise of the input differential pair dominates for both single-stage and multi-stage OTA. $S_{eq}(f)$ can be expressed as $4kT/\alpha g_{m1}$, where $g_{m1}$ is
the transconductance of the input differential pair and the excess noise factor $\alpha$ accounts for noise contributions from the devices other than the input differential pair, as well as any potential excess thermal noise in the input devices themselves.

The noise transfer function (NTF) from $v_{n,eq}$ to $v_{n,out}$ is given by

$$v_{n,out}/v_{n,eq} = \frac{1}{\beta} \cdot \frac{1}{1 + j2\pi f/\omega_p}$$

(4.6)

where the pole frequency $\omega_p$ of the NTF is the same as that of STF.

The total output noise power can be calculated by integrating from dc to infinite frequency

$$\overline{v_{n, out}^2} = \int_0^{\infty} S_{eq}(f) \cdot \left(\frac{1}{\beta}\right)^2 \cdot \left|\frac{1}{1 + j2\pi f/\omega_p}\right|^2 df$$

$$= S_{eq}(f) \cdot \left(\frac{1}{\beta}\right)^2 \cdot \omega_p = S_{eq}(f) \cdot \left(\frac{1}{\beta}\right)^2 \cdot \frac{\pi f_{3dB}}{2}.$$  

(4.7)

Equations (4.5) and (4.7) show that both settling accuracy and noise power are proportional to the THA bandwidth.

### 4.2.2 Switched-bandwidth THA

The signal and the noise have different time-dependence during the amplification phase. The output signal takes the entire amplification phase to settle, whereas the output noise mostly depends on the bandwidth at the end of the phase. Hence, the THA bandwidth can be varied during the settling time to optimize the noise. For example, the THA starts with a high bandwidth and settles close to the final voltage value at $T_s/2$. Then it changes to a lower bandwidth and continues settling. While achieving the same settling accuracy, a much lower output noise is obtained due to the lower bandwidth. This technique is introduced as a switched-bandwidth amplifier.
CHAPTER 4. NOISE REDUCTION TECHNIQUE THROUGH BANDWIDTH SWITCHING

Figure 4.5: Transient response of the switched-bandwidth THA and the conventional THA with \( \omega_p = 6.93/T_s, \omega_{p1} = 1.5\omega_p, \omega_{p2} = 0.5\omega_p \) and \( T_1 = T_2 = 0.5T_s \).

The time allocation of the amplification phase is shown in Figure 4.4. The THA output is sampled by the following SC circuit at \( t_s \). The amplification phase is split into two consecutive sub-phases, \( \phi_4 \) and \( \phi_5 \), lasting \( T_1 \) and \( T_2 \) respectively. The associated pole is \( \omega_{p1} \) and \( \omega_{p2} \) in each sub-phase with the bandwidth switched at \( t_m \). Obviously, \( \omega_{p2} \) should be less than \( \omega_{p1} \) to reduce noise.

The THA becomes a piecewise single-pole system. The transient response of \( V_{out} \) can be expressed as

\[
V_{out}(t) = \begin{cases} 
V_s \cdot (1 - e^{-\omega_{p1}t}), & 0 \leq t < t_m \\
V_s \cdot (1 - e^{-\omega_{p1}T_1} \cdot e^{-\omega_{p2}(t-T_1)}), & t_m \leq t < t_s.
\end{cases}
\]

The settling error after \( T_s \) is \( e^{-\left(T_1\omega_{p1} + T_2\omega_{p2}\right)} \) according to (4.8). To achieve the same settling accuracy as the conventional THA with constant pole \( \omega_p \), the requirements of \( \omega_{p1}, \omega_{p2}, T_1 \) and \( T_2 \) are

\[
T_1 + T_2 = T_s, \quad (4.9)
\]

\[
\omega_{p1}T_1 + \omega_{p2}T_2 = \omega_pT_s. \quad (4.10)
\]

Therefore, the settling error of the switched-bandwidth THA only depends on the time-averaged bandwidth in the amplification phase.
CHAPTER 4. NOISE REDUCTION TECHNIQUE THROUGH BANDWIDTH SWITCHING

Figure 4.5 shows an example of the transient responses for 10-bit settling accuracy. The pole of a constant bandwidth THA is \( \omega_p = 6.93/T_s \). A switched-bandwidth amplifier with \( \omega_{p1} = 1.5\omega_p \) and \( \omega_{p2} = 0.5\omega_p \) and \( T_1 = T_2 = 0.5T_s \) settles to the same accuracy.

The noise of the switched-bandwidth amplifier can be analyzed in a piecewise manner as well. The noise PSD in \( \phi_4 \) and \( \phi_5 \) is \( S_{eq,1}(f) \) and \( S_{eq,2}(f) \), respectively. In \( \phi_4 \), from 0 to \( t_m \), the output noise can be considered as stationary due to high bandwidth \( \omega_{p1} \). The noise power at \( t_m \) is derived in a similar way to (4.7),

\[
\frac{v_{n,\text{out},s}^2(t_m)}{v_{n,\text{out},s}^2(0)} = S_{eq,1}(f) \cdot \left( \frac{1}{\beta} \right)^2 \cdot \frac{\omega_{p1}}{4}.
\]  

(4.11)

During \( \phi_5 \), since the assumption that the output noise is stationary may be invalid due to smaller \( \omega_{p2} \), the output noise power no longer follows (4.7). Resorting to the first-order stochastic differential equation (Langevin equation [42]), the output noise power is given by

\[
\frac{v_{n,\text{out},s}^2(t)}{v_{n,\text{out},s}^2(t_m)} = e^{-2(t-T_1)\omega_{p2}} + S_{eq,2}(f) \cdot \left( \frac{1}{\beta} \right)^2 \cdot \frac{\omega_{p2}}{4} \cdot \left( 1 - e^{-2(t-T_1)\omega_{p2}} \right).
\]  

(4.12)

The first term on the right hand side of (4.12) is the decaying noise power inherited from \( \phi_4 \), whereas the second term is generated during \( \phi_5 \).

Evaluating (4.12) at \( t = t_s \) and substituting (4.11) in (4.12) leads to

\[
\frac{v_{n,\text{out},s}^2(t_s)}{v_{n,\text{out},s}^2(t_m)} = S_{eq,1}(f) \cdot \left( \frac{1}{\beta} \right)^2 \cdot \frac{\omega_{p1}}{4} \cdot e^{-2T_2\omega_{p2}} + S_{eq,2}(f) \cdot \left( \frac{1}{\beta} \right)^2 \cdot \frac{\omega_{p2}}{4} \cdot \left( 1 - e^{-2T_2\omega_{p2}} \right).
\]  

(4.13)

For a single-stage OTA, \( G_m \) is equal to the transconductance of the input differential pair and appears to be the only choice for changing bandwidth as shown in (4.2). However, the product of the noise PSD and bandwidth is constant regardless of the bandwidth, and hence no noise reduction benefit is achieved compared to the conventional THA.

In contrast, an OTA is typically implemented as a cascade of two or more stages to provide adequate gain for 10 bit or higher applications. Multi-stage architectures offer a variety of ways to switch bandwidth while keeping input-referred noise PSD relatively constant. For example, the Miller capacitance can be adjusted for a two-stage Miller amplifier.

The proposed OTA in this chapter includes a pre-amplifier with gain \( A_1 \) followed by a transconductance stage \( g_{m2} \), with \( G_m \) equal to \( A_1 g_{m2} \). Either \( A_1 \) or \( g_{m2} \) can be adjusted to control the bandwidth. The proposed OTA changes \( g_{m2} \) and keeps the pre-amplifier untouched to obtain the additional slew rate benefit to be described in Section 4.3. The noise contribution of \( g_{m2} \) stage is negligible due to the pre-amplifier gain, and hence the input-referred noise PSD can be considered approximately the same as that of the constant bandwidth amplifier during both sub-phases.
In this case, the ratio of the output noise power of the switched-bandwidth THA to that of the constant bandwidth THA is given by

\[ Q_n(\omega_p1, \omega_p2, T_1, T_2) = \frac{v_{n, out, s}(t_s)}{v_{n, out}^2} \approx \frac{\omega_p1}{\omega_p} \cdot e^{-2T_2\omega_p2} + \frac{\omega_p2}{\omega_p} \cdot (1 - e^{-2T_2\omega_p2}) \].

(4.14)

For three combinations of \( T_1 \) and \( T_2 \), \( Q_n \) is plotted in Figure 4.6 as a function of \( \omega_p2/\omega_p \), with \( \omega_p1, \omega_p2, T_1 \) and \( T_2 \) satisfying (4.9) and (4.10).

As shown in Figure 4.6, \( Q_n \) indicates the noise reduction of the switched-bandwidth THA. \( Q_n \) can be reduced by lowering \( \omega_p2/\omega_p \) until hitting an optimum point, below which \( Q_n \) rises rapidly and exceeds 1 when \( \omega_p2/\omega_p \) is close to 0. A detailed explanation is as follows. When \( \omega_p2 \) is small, \( \omega_p1 \) has to be large to meet the same settling accuracy according to (4.10). As a result, the noise power generated in \( \phi_4 \) becomes larger and decays more slowly in \( \phi_5 \) according to (4.13). In the extreme case where \( \omega_p2 = 0 \) and \( T_1 = T_2 = 0.5T_s \), \( \omega_p1 \) must be \( 2\omega_p \) to satisfy (4.10). The output settles to the desired accuracy by the end of \( \phi_4 \), but with twice the noise power, and then it is “frozen” during \( \phi_5 \) due to zero bandwidth.

In Figure 4.6, the minimum \( Q_n = 0.42 \) is obtained at \( \omega_p2 = 0.33\omega_p, \omega_p1 = 3\omega_p, T_1 = 0.25T_s \) and \( T_2 = 0.75T_s \). Notably the optimum region is shallow and not sensitive to the combination of \( T_1 \) and \( T_2 \). For example, with \( T_1 = T_2 = 0.5T_s \), the minimum \( Q_n \) is 0.46. Therefore,
CHAPTER 4. NOISE REDUCTION TECHNIQUE THROUGH BANDWIDTH SWITCHING

Figure 4.7: Schematic of a conventional two-stage OTA.

Figure 4.8: Cascode NMOS in the transconductance stage of OTA. (a) schematic. (b) transient response of internal nodes.

The switched-bandwidth technique is robust to circuit parameter variations. In addition, a moderate $\omega_{p2}$ is good for dissipating any output glitch caused by switching bandwidth.

The power consumption of the pre-amplifier is the same for both the constant bandwidth and switched-bandwidth amplifier. The power consumption of the transconductance stage is given by

$$P_{gm2} = V_{dd} \cdot I_{D2} = V_{dd} \cdot \left( \frac{g_{m2}}{I_{D2}} \right)^{-1} \cdot g_{m2}$$

(4.15)

where $g_{m2}/I_{D2}$ is set by the current density and is designed to be constant. Hence, the instantaneous
Chapter 4. Noise Reduction Technique through Bandwidth Switching

$P_{gm2}$ is proportional to $g_{m2}$, and thus is proportional to the THA bandwidth at any moment. The average $P_{gm2}$ is proportional to the average bandwidth, i.e. $(\omega_{p1}T_1 + \omega_{p2}T_2)/T_s$, which is the same as the conventional constant bandwidth amplifier based on (4.10). The switched-bandwidth technique reallocates the power consumption between two sub-phases.

Overall, the above analysis is based on a single-pole model and shows that the switched-bandwidth technique can reduce the output noise power by more than 50% with the same settling accuracy and no extra power consumption. Although noise contributions from switches are not included in the analysis, they are also reduced because the NTFs of switches have the same pole as that of the OTA.

Compared to the incomplete settling technique in [42], the proposed switched-bandwidth technique...
amplifier has its output fully settled to the desired accuracy, eliminating the need for calibration and significantly reducing the design complexity. In addition, the proposed technique is less sensitive to clock jitter than both the incomplete settling technique and the conventional amplifier, due to smaller \( \frac{dV_{out}}{dt} \) at \( t_s \).

### 4.3 Slew Rate Improvement via Cascode Bias Switching

A conventional fully differential OTA with wide-band pre-amplifier [52] is shown in Figure 4.7. The resistor-loaded pre-amplifier has a moderate gain \( A_1 = g_{m1}R \), where \( g_{m1} \) is the transconductance of \( M_{1p} \) and \( M_{1n} \), and \( R \) is the resistor load. The transconductance stage \( g_{m2} \) is a cascode push-pull stage to provide high DC gain. The dominant and non-dominant poles are at the OTA outputs and the pre-amplifier outputs, respectively. This OTA suffers slew rate degradation due to the cascode output stage. The proposed technique below improves the slew rate by switching the
cascode bias voltage.

The swing of each single-ended output is from $V_{\text{min}}$ to $V_{\text{max}}$, while the differential swing is from $-V_s$ to $V_s$, where $V_s = V_{\text{max}} - V_{\text{min}}$. The following analysis focuses on the single-ended cascode NMOS transistors of the transconductance stage as shown in Figure 4.8a. Similar analysis can be applied to the cascode PMOS transistors. In order to achieve high DC gain, both $M_2$ and $M_3$ should be in the saturation region at the end of the amplification phase. The cascode bias voltage $V_{bn}$ is designed to force the drain source voltage of $M_2$ ($V_{DS2}$) equal to its overdrive voltage ($V_{ov2}$) plus additional headroom.

At the beginning of the amplification phase, a step $V_s/c$ applied at the THA input triggers the transient response at the OTA internal nodes as shown in solid lines of Figure 4.8b causing the output $V_{outp}$ to settle to $V_{\text{min}}$ eventually. The node $V_1$ sees a step $V_{s1}$ and the instantaneous gate source voltage of $M_2$ is increased by $V_{s1}$. The dip on the node $V_X$ ($V_{sX}$) is roughly $V_{s1} \cdot (g_{m,M2}/g_{m,M3})$ for constant $V_{bn}$, where $g_{m,M2}$ and $g_{m,M3}$ are the transconductance of $M_2$ and $M_3$, respectively. The instantaneous drain source voltage of $M_2$ is reduced by $V_{sX}$, which is likely to be larger than the headroom of $M_2$, driving $M_2$ into the triode region and deteriorating the slew rate. The instantaneous drain current of $M_2$ during the transient is given by

$$I_{D2,c} = K \left( (V_{ov2} + V_{s1}) (V_{DS2} - V_{sX}) - \frac{(V_{DS2} - V_{sX})^2}{2} \right)$$

$$\approx K (V_{ov2} + V_{s1}) (V_{DS2} - V_{sX})$$

(4.16)

where $K$ is the current factor of $M_2$.

The $G_m$ of OTA has to increase in order to compensate for the slew rate degradation. As a result, the OTA noise power will be larger than the value predicted by the single-pole model due to the increase of the small-signal bandwidth.

However, $V_{bn}$ only needs to keep $M_2$ and $M_3$ in the saturation region when the output approaches its final value. The amplification phase can be split into two sub-phases as well. In the first sub-phase, $V_{bn}$ is set to be high to increase the voltage of $V_X$ and keep $M_2$ in the saturation region, enhancing the slew rate. $V_{bn}$ is then reduced to achieve high DC gain during the second sub-phase. The transient response of this technique is shown in dashed lines of Figure 4.8b. The instantaneous drain current of $M_2$ is given by

$$I_{D2,s} = \frac{K}{2} (V_{ov2} + V_{s1})^2.$$  

(4.17)
The slew rate improvement \( Q_{SR} \), defined as the ratio of the instantaneous \( M_2 \) current using this technique to that of the conventional cascode output stage, is given by

\[
Q_{SR} = \frac{I_{D2,s}}{I_{D2,c}} \approx \frac{V_{ov2} + V_{s1}}{2(V_{DS2} - V_{SX})}.
\]

(4.18)

Note that switching \( V_{bn} \) every clock cycle requires strong driving capability, which can induce a significant amount of power consumption. In Section 4.4, it will be shown that the proposed circuit achieves the desired transient response without incurring a significant power penalty.

Unlike the amplifier in [50], which increases the slew rate based on the input magnitude, the proposed technique does not require any knowledge of the input, which simplifies the design and reduces the parasitic capacitance at the OTA inputs.

### 4.4 Circuit Details and Simulation Results

Combining both techniques described in Section 4.2 and Section 4.3 leads to the proposed OTA shown in Figure 4.9a. As compared to the conventional OTA, the transconductance stage of the proposed OTA is split into two branches. The size ratio of \( M_{21} \) to \( M_{22} \) and of \( M_{31} \) to \( M_{32} \) is 3-to-4. \( M_{21} \) always connects to cascode \( M_{31} \), whereas \( M_{22} \) connects to a separate cascode \( M_{32} \) through the switch \( S_n \). PMOS transistors have a similar configuration.

The OTA can operate in two modes - constant mode and switched mode. The reconfigurability enables a performance comparison using the same chip. In the constant mode, \( V_{bn2} \) is equal to \( V_{bn} \) and the switches \( S_n \) and \( S_p \) are always on, which is equivalent to a conventional OTA. In the switched mode, \( V_{bn2} \) is 300 mV higher than \( V_{bn} \) and the switches \( S_n \) and \( S_p \) are controlled by \( \phi_4 \) as shown in Figure 4.9b. During \( \phi_4 \), both \( M_{21} \) and \( M_{22} \) are available to sink current from the output. \( M_{21} \) has the same sink capability in both modes, while the instantaneous current sink of \( M_{22} \) is much larger in the switched mode because \( M_{22} \) remains saturated. The average slew rate improvement is

\[
Q_{SR,s} = \frac{4 \times Q_{SR} + 3}{7}.
\]

(4.19)

During \( \phi_5 \), \( S_n \) and \( M_{22} \) are turned off and \( V_{out} \) settles to its final value with reduced bandwidth, resulting in the noise benefit described in Section 4.2. In addition, note that \( V_{bn} \) does not switch every clock cycle. Instead, both high and low static bias voltages are provided to the separate cascodes and the signal path through \( V_{bn2} \) is switched on and off by \( S_n \).

The THA design specs are 12 dB closed-loop gain and 0.1\% or better non-linearity at 90 MS/s. The differential output swing is from \(-800\) mV to \(800\) mV, and the single-ended output

51
swing is from 400 mV to 1.2 V. The amplifier drives a total 7.5 pF capacitive load. The closed-loop bandwidth in the switched mode is 78.5 MHz at the end of the amplification phase, while in the constant mode it is 185 MHz. The bandwidth ratio is 43%, which matches the 3-to-7 device size ratio of M\textsubscript{21} to the sum of M\textsubscript{21} and M\textsubscript{22}. The pre-amplifier of the OTA has a gain of 16 dB.

Figure 4.10 shows the transient simulation results with the constant mode in solid lines and the switched mode in dashed lines. The top and bottom subplots show the transient waveforms of the differential output and the drain of M\textsubscript{22} (V\textsubscript{X2}), respectively. The differential output settles to the worst case −800 mV. When operating at 90 MS/s with 50% duty cycle, each clock period is 11.1 ns and the total settling time is 5 ns after deducting the non-overlapping periods. When the THA operates in the constant mode, the overdrive voltage of M\textsubscript{21} and M\textsubscript{22} is 100 mV and the headroom voltage is 50 mV. Both V\textsubscript{X1} and V\textsubscript{X2} drop from 150 mV to 50 mV and drive M\textsubscript{21} and M\textsubscript{22} into the triode region. In contrast, V\textsubscript{X2} remains above 200 mV and M\textsubscript{22} stays in the saturation region in the switched mode. The simulated improvement of the slew rate $Q_{sr,s}$ is 1.6, while the calculated $Q_{sr,s}$ is 1.86 according to (4.19). The calculation is a rough estimate, but still matches reasonably well with the simulation result. While the settling is non-linear due to the non-dominant pole effect and slew rate limitations, the time for settling to the final value ±0.05% is 4.56 ns in the switched mode, 0.38 ns faster than the constant mode.

The noise power ($P_{n,a}$) is simulated using a Periodic Noise analysis and plotted in Figure 4.11a as a function of time in the amplification phase. $Q_n$, defined as the ratio of $P_{n,a}$ in switched mode ($P_{n,a,switch}$) to $P_{n,a}$ in constant mode ($P_{n,a,const}$), is plotted in Figure 4.11b, which shows clearly that $P_{n,a,switch}$ is lower during $\phi_5$ due to decreased THA bandwidth. At 5 ns, the instant when the following SC stage samples the THA output, $P_{n,a,const}$ is $2.83 \times 10^{-8} V^2$ and $P_{n,a,switch}$ is $1.50 \times 10^{-8} V^2$. $Q_n$ is 53% and is slightly larger than the bandwidth ratio (43%). This discrepancy is caused by two mechanisms. The first mechanism is the residual noise from $\phi_4$ as described by the first term of (4.13), which would further decay with more settling time as shown in Figure 4.11b. The other mechanism is due to the extra noise contribution of the transconductance stage in the switched mode. Smaller transconductance causes slightly larger input-referred noise PSD.

From the design perspective, the size ratio of M\textsubscript{21} to M\textsubscript{22} and the duration of $\phi_4$ are critical to determine the THA settling accuracy and noise. M\textsubscript{21} is comparable to M\textsubscript{22}, which results in optimal noise performance as shown in Figure 4.6. Simulations are used to determine the optimum combination of these two parameters. $\phi_4$ is generated by a timer circuit which is not discussed in this chapter. The simulated duration of $\phi_4$ is 2.7 ns ±5% across the design operating range.
Figure 4.11: Simulated $P_{n,a}$ of the proposed THA vs. time in amplification phase. (a) $P_{n,a,\text{const}}$ and $P_{n,a,\text{switch}}$. (b) ratio of $P_{n,a,\text{switch}}$ to $P_{n,a,\text{const}}$.

4.5 Experimental Results

A THA followed by a 14-bit ADC forms a complete signal channel. Four identical channels are implemented on the same chip. The test chip is fabricated in a 65 nm standard CMOS process.
The die photograph is shown in Figure 4.12.

The ADC linearity is designed and measured to be 14-bit accurate. Therefore, the INL of the signal channel reflects the non-linearity of the THA, which is mainly caused by settling error. The measured INL of each channel is shown in Figure 4.13a and 4.13b and the channel-to-channel variation is negligible. The peak-peak INL of each channel is listed in Figure 4.13c. All of the four THAs achieve the target of 0.1% non-linearity at 90 MS/s in both modes, and the INL in the switched mode is improved, which agrees with simulation results predicting improved settling.

The inputs of the four signal channels are grounded and one million ADC output samples are obtained for each channel. The standard deviation of the samples, referred to the THA output, indicates the total channel noise power ($P_n$) \[^{16}\] and is composed of three parts: the THA tracking phase noise ($P_{n,t}$), the THA amplification phase noise ($P_{n,a}$), and the ADC noise ($P_{n,\text{ADC}}$). The
measured standard deviations (noise rms) of four channels are listed in Table 4.1. It is clear that the noise power in the switched mode is lower than that in the constant mode for each channel. Figure 4.14 shows the measured histograms of one of the channels.

\( P_{n,t} \) and \( P_{n,ADC} \) need to be extracted from the total channel noise power in order to derive \( P_{n,a} \). In the constant mode, the OTA bandwidth can be reduced by changing \( g_{m2} \). These bandwidth settings only affect \( P_{n,a} \). \( P_{n,a,max} \) is the noise power with maximum bandwidth, and \( k_{BW} \) is defined as the ratio with \( P_{n,a} \) of the reduced bandwidth setting to \( P_{n,a,max} \). Thus, the total channel noise power \( P_{n,const}(BW) \) satisfies

\[
P_{n,const}(BW) = P_{n,t} + P_{n,ADC} + P_{n,a,max} \cdot k_{BW}. \tag{4.20}
\]

The simulated \( k_{BW} \) for each bandwidth setting is shown in Table 4.2. Since \( k_{BW} \) is the ratio of the noise power, the simulated value should match the measurement results reasonably well.
CHAPTER 4. NOISE REDUCTION TECHNIQUE THROUGH BANDWIDTH SWITCHING

![Graph](image)

Figure 4.14: Measured histograms of one channel output, with mean value removed.

Table 4.1: Measured standard deviation of four channel outputs referred to THA output

<table>
<thead>
<tr>
<th>Channel #</th>
<th>Standard deviation (µV-rms)</th>
<th>Constant mode</th>
<th>Switched mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>320.4</td>
<td>308.9</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>324.1</td>
<td>299.7</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>317.1</td>
<td>286.7</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>321.4</td>
<td>299.5</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: Simulated ratio of $P_{n,a}$ of the reduced bandwidth setting to $P_{n,a,max}$

<table>
<thead>
<tr>
<th>BW setting</th>
<th>$k_{BW}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>100%</td>
</tr>
<tr>
<td>A</td>
<td>62%</td>
</tr>
<tr>
<td>B</td>
<td>48%</td>
</tr>
<tr>
<td>C</td>
<td>35%</td>
</tr>
</tbody>
</table>

The measured average $P_{n,\text{const}(BW)}$ of four channels versus $k_{BW}$ is plotted in Figure 4.15. The four data points fit a straight line, which is consistent with the previous analysis. The y-axis intercept is extrapolated to obtain the sum of $P_{n,t}$ and $P_{n,\text{ADC}}$ as $7.25 \times 10^{-8} \text{ V}^2$.

The measured average $P_n$ of the four channels in the constant mode (maximum bandwidth) and the switched mode is $10.29 \times 10^{-8} \text{ V}^2$ and $8.92 \times 10^{-8} \text{ V}^2$, respectively. After subtracting the sum of $P_{n,t}$ and $P_{n,\text{ADC}}$, $P_{n,a}$ in the constant mode and the switched mode is $3.04 \times 10^{-8} \text{ V}^2$ and $1.67 \times 10^{-8} \text{ V}^2$, respectively. Their ratio $Q_n$ is 55%. Compared with the simulation results in
Section 4.4, the absolute value of \( P_{n,a} \) in both modes is about 10% higher, which may be caused by inaccurate noise model or incorrect parasitic capacitance estimate. However, the relative reduction is predicted accurately by simulation.

The total channel noise in Table 4.1 does not show dramatic improvement because \( P_{n,a} \) is not the dominant component of this test chip. However, for implementation in which \( P_{n,a} \) dominates, this technique could achieve larger noise reduction.

The THA consumes 15.6 mW in the constant mode and 14.7 mW in the switched mode. The power consumption reduction results from the fact that \( 4/7 \) of the transconductance stage is turned off for almost half of the amplification phase.

In summary, the proposed THA reduces the amplification phase noise power by nearly half as compared with the conventional THA. Improved linearity and lower power consumption are accomplished simultaneously, while requiring one extra clock phase and slightly more complex circuit design.
CHAPTER 4. NOISE REDUCTION TECHNIQUE THROUGH BANDWIDTH SWITCHING

4.6 Conclusion

Traditionally, switched-capacitor amplifiers face tradeoffs among power consumption, noise, and accuracy. This chapter presents a novel amplifier that allows simultaneous improvements in all three dimensions by switching the bandwidth. The amplification phase is split into two sub-phases. In the first sub-phase, the amplifier is configured with high bandwidth and high slew rate. In the second sub-phase, the bandwidth is reduced, achieving the target settling accuracy but with much lower noise. Test chip measurements demonstrate that the proposed THA achieves 45% noise power reduction in the amplification phase as well as improved linearity and lower power consumption. The proposed techniques provide a new degree of freedom in design for low-noise and low-power SC circuits.
Chapter 5

Current Source Mismatch and Calibration Techniques

5.1 Current Mismatch and Temperature Dependence

Current-steering DACs are widely used to generate analog signals across a wide speed and resolution range. High accuracy applications require the current sources of a DAC to match accurately. The mismatches between the current sources directly affect the DAC’s static linearity. The random device mismatch during manufacture causes the current source mismatch.

A current source in modern CMOS processes is normally designed with transistors much larger than the minimum feature size. Its current can be described by the drain current in saturation,

\[ I_{cs} = \frac{\mu C_{ox} W}{2} \left( V_{GS} - V_{TH} \right)^2 = \frac{\beta}{2} (V_{GS} - V_{TH})^2. \] (5.1)

Two current sources consisting of identically sized and biased transistors are shown in Figure 5.1. The random mismatch between two transistors can be decomposed into two components, i.e. the threshold voltage mismatch \( \Delta V_{TH} \) and the current factor mismatch \( \Delta \beta / \beta \) [23]. \( \Delta V_{TH} \) and \( \Delta \beta / \beta \) are unknown during design, but fixed after fabrication. Hence, the mismatch between \( I_{cs} \) and the reference current \( I_{ref} \) is given by

\[ \Delta I = \frac{\Delta \beta}{\beta} I + g_m (-\Delta V_{TH}), \] (5.2)

where \( I \) is the nominal bias current and \( g_m = \sqrt{2\mu C_{ox}(W/L)I} \) is the nominal transconductance.

\( I \) is usually designed to be temperature independent, whereas \( g_m \) shows strong dependence on temperature due to the mobility \( \mu \). The simulated \( g_m \) of a current source in 65 nm CMOS
process increases by 50% as the temperature changes from 120 °C to −40 °C, as shown in Figure 5.2. Therefore, the current mismatch is a strong function of the temperature $T$ and is given by

$$\Delta I(T) = \frac{\Delta \beta}{\beta} I + g_m(T) (-\Delta V_{TH}).$$

The temperature dependence exacerbates the mismatch problem.

The normalized current mismatch is

$$\frac{\Delta I(T)}{I} = \frac{\Delta \beta}{\beta} + \frac{g_m(T) (-\Delta V_{TH})}{I}.$$
CHAPTER 5. CURRENT SOURCE MISMATCH AND CALIBRATION TECHNIQUES

Figure 5.3 shows Monte Carlo simulation results versus temperature with each line representing an individual Monte Carlo trial. $\Delta V_{TH}$ and $\Delta \beta / \beta$ determine the slope and offset for each current source, respectively. Because of decreasing $g_m$, $|\Delta I| / I$ decreases with increasing temperature.

## 5.2 Conventional Calibration Techniques

A straightforward method to reduce the device mismatch is to increase the device physical area. Quadrupling the transistor physical area typically reduces mismatch by 1 bit. However, sizing the current source transistors to yield 16-bit matching requirements is impractical and furthermore leads to large parasitic capacitance, which limits the dynamic linearity at high frequency. Calibration techniques can improve the matching of current sources of small area, enhancing both static and dynamic linearity.

Background calibration techniques [53, 54, 55, 56, 57, 58] continuously calibrate the current sources during normal operation, hence tracking the operating temperature changes. However, they consume extra power due to the active calibration circuits such as sensing and controlling circuits, and tend to degrade the performance by introducing calibration activity noise and spurs.

Foreground calibration techniques [59, 60, 24] are applied when the DAC is inactive such as during power up, stand by, or test phase, hence avoid the disadvantages of background calibration.
An auxiliary DAC, termed CAL DAC, provides a small correction current to compensate the current source mismatch. The CAL DAC input code stores the mismatch information in digital format. However, foreground calibration loses its accuracy when the operating environment (e.g. temperature) deviates from the calibration environment. This disadvantage stems from the different temperature dependence of the CAL DAC compared to that of the current source mismatch. To further increase the output impedance and improve the dynamic linearity for GHz DACs, a small current array is utilized, which exacerbates the temperature variations of the current mismatch and necessitates the need of a temperature insensitive calibration technique.

Some other calibration techniques are implemented in recent publications \[61, 62\] to improve the linearity by reassigning the switching sequence based on the measurement of the current sources instead of using the CAL DACs. However, there is no discussion about their temperature stability.

Figure 5.4 shows the schematic view of the calibrated current source (CCS) in \[24\], which is composed of the main current source and a CAL DAC. The reference current, $I_{\text{ref}}$, is a replica of CCS with the CAL DAC input code fixed at mid-scale. Hence, the CAL DAC can provide either positive or negative correction current (denoted as $I_{\text{cor}}$). The current of the CCS is

$$I_{\text{ccs}} = I_{\text{cs}} + I_{\text{cor}} = I_{\text{ref}} + \Delta I + I_{\text{cor}}. \quad (5.5)$$

The calibration configuration is shown in Figure 5.5. During calibration, the cascode switch $M_{\text{CAS}1}$ is off and $M_{\text{CAS}2}$ is on, thereby steering the current of the CCS to the calibration.
circuitry that includes a comparator and a calibration control circuit. All the CCSs are calibrated against the on-chip reference current $I_{ref}$. Their difference is adjusted to approach zero through the successive approximation register (SAR) logic.

Based on the previous discussion, it is easy to understand why the conventional calibration technique in [24] does not track the temperature. The CCS schematic proposed in [24] is shown in Figure 5.6. $M_{cal}$ provides the full scale current (FSC) for the CAL DAC, which is proportional to $I$ because $M_{CS}$ and $M_{cal}$ share the same gate voltage. The transistors from $M_{s0}$ to $M_{sn}$ forms a 6-bit current splitter. The CAL DAC output current $I_{cor}$ is controlled by the digital inputs of the switches from $S_1$ to $S_n$. After the CCS is calibrated at a certain temperature $T_{cal}$, $I_{cor}$ is equal to $-\Delta I(T_{cal})$ at $T_{cal}$. When the temperature varies, $-\Delta I(T)$ will change, whereas $I_{cor}$ still remains the same. Therefore, the CAL DAC does not track the current mismatch and the linearity of the DAC degrades with the varying temperature.
Figure 5.6: Schematic of CCS in [Mercer JSSC 2007].
Chapter 6

DAC Calibration Technique Tracking Temperature Variations

6.1 Introduction

There has been extensive research on the calibration techniques, but limited research on the capability to track temperature variations. A temperature insensitive DAC calibration is proposed in [63] by making use of redundant current source units at the cost of much more hardware complexity. Another disadvantage is its incapability to control the calibration accuracy.

In this chapter, a foreground DAC calibration technique named \( g_m \) calibration is proposed to track temperature variations. Similar to [24], the CAL DACs are implemented to provide the correction current for each current source, but the correction current tracks the current source mismatch change with temperature. Therefore, better DAC linearity over a wide temperature range can be achieved.

This chapter is organized as follows. Section 6.2 presents the \( g_m \) calibration technique and its implementation in a 14-bit DAC. Section 6.3 presents the transistor level simulation results, followed by the conclusions in Section 6.4.
6.2 $g_m$ Calibration Technique

6.2.1 Basic Idea

The mismatch between $I_{cs}$ and the reference current $I_{ref}$ is given by (5.2). In a typically biased current source, $|g_m \Delta V_{TH}|$ is usually larger than $|(\Delta \beta / \beta)I|$ and dominates the total current mismatch. Therefore, the current mismatch can be expressed as

$$\Delta I \approx -g_m \Delta V_{TH}. \quad (6.1)$$

The temperature dependence of $\Delta V_{TH}$ is measured in [64]. The conclusion is that “the standard deviation of threshold-voltage mismatch is practically constant” with respect to temperature. $\Delta V_{TH}$ is considered as constant in this work. Therefore, the temperature dependence of $g_m$ determines the temperature dependence of the current mismatch.

The schematics of the CCS and the CAL DAC in the $g_m$ calibration technique are shown in Figure 6.1 and Figure 6.2, respectively. The difference comparing with the conventional calibration technique [24] is that the FSC of the CAL DAC is proportional to $g_m V_{ref}$, where $V_{ref}$ is a constant voltage derived from the bandgap voltage. When the current source is calibrated at $T_{cal}$, $(g_m(T_{cal}) \Delta V_{TH})/(g_m(T_{cal}) V_{ref})$, i.e. $(\Delta V_{TH} / V_{ref})$, is captured and stored as the CAL DAC digital value. When the temperature drifts to $T$, the CAL DAC output current $I_{cor}$ is equal to $g_m(T) \Delta V_{TH}$, which compensates the current mismatch at the new temperature $T$. 

Figure 6.1: Calibrated Current Source (CCS) of $g_m$ calibration technique.
6.2.2 DAC Configuration

A 14-bit DAC is designed to verify the proposed $g_m$ calibration technique. The DAC segmentation is 4-4-6, i.e. 4 most significant bits (MSBs), 4 intermediate significant bits (ISBs) and 6 least significant bits (LSBs). The DAC current sources are implemented using PMOS transistors. The MSB segment is thermometer coded and consists of 15 MSB current sources $M_{CS,MSB}$. There are 16 ISB current sources $M_{CS,ISB}$. 15 of them belong to the thermometer coded ISB segment. The 16-th $M_{CS,ISB}$ is subdivided to provide the remaining 6 binary coded LSBs using a current splitter. The width of $M_{CS,MSB}$ is 16 times as large as the width of $M_{CS,ISB}$. The DAC reference current is 40 mA. Each $M_{CS,MSB}$ is 2.5 mA and each $M_{CS,ISB}$ is 156.25 µA. The 14-bit level LSB is 2.44 µA.

The calibration configuration is the same as shown in Figure 5.5. Each current source is connected with a new CAL DAC to form a CCS. Each CAL DAC is 6-bit and the resolution is 1/8 14-bit level DAC LSB. Based on the foundry measurement results, the standard deviation of the random mismatch of $M_{CS,MSB}$ is 1.94 µA. The FSC of the CAL DAC is 19.52 µA, large enough to cover 6σ range. The current splitter of the LSB segment relies on its intrinsic mismatch. All 16 ISB CCSs are calibrated to the reference current $I_{ref}$ first. Then, the sum of all 16 ISB CCSs is used as the reference current in order to calibrate 15 MSB CCSs to eliminate the boundary from ISB to MSB.
6.2.3 $g_m V_{\text{ref}}$ Current Generator

A $g_m V_{\text{ref}}$ current generator, shown in Figure 6.3, is needed for all the CAL DACs. There is a sub-circuit named “$I_R$ generator” inside the $g_m V_{\text{ref}}$ current generator. The voltage across $R_I$ is forced to the bandgap voltage ($V_{BG}$) by the amplifier $A_1$. The current ($I_R$) is equal to $V_{BG}/R_I$. Its mirror copy is fed into the two resistors, $R_1$ and $R_2$, with the same value $R$. $R_I$, $R_1$ and $R_2$ are the same type of on-chip resistors. The voltage across $R_1$ and $R_2$ is $(V_{BG}R)/(R_I)$, which is a constant fraction of the bandgap voltage. $V_{GS}$ is the gate voltage of the current sources generated by the DAC bias circuit. $M_1$ and $M_2$ are replicas of $M_{CS, MSB}$ with the same $g_m$. Therefore, the gate voltages of the differential pair $M_1$ and $M_2$ are $V_{GS} + I_R R$ and $V_{GS} - I_R R$, respectively.

The output current, $I_{gm}$, is equal to the current difference of $M_1$ and $M_2$,

$$I_{gm} = \frac{\beta}{2} (V_{GS} + I_R R - V_{TH})^2 - \frac{\beta}{2} (V_{GS} - I_R R - V_{TH})^2$$

$$= g_m \times 2I_R R. \quad (6.2)$$

By defining $V_{\text{ref}}$ to be $2I_R R = (2V_{BG}R)/(R_I)$, the output current $I_{gm}$ becomes $g_m V_{\text{ref}}$. 

Figure 6.3: Schematic of $g_m V_{\text{ref}}$ current generator.
CHAPTER 6. DAC CALIBRATION TECHNIQUE TRACKING TEMPERATURE VARIATIONS

6.3 Transistor Level Simulations

The DAC is implemented in a 65 nm CMOS process and simulated using BSIM4 models. Random mismatches are attached to each transistor with the parameters provided by the foundry. $\Delta \beta/\beta$ is equally divided into the mismatch of width, length and gate oxide thickness. $\Delta V_{TH}$ is modeled in the parameter $V_{TH0}$.

Three scenarios are simulated for illustration. In the first scenario, there is no calibration and the DAC purely relies on the intrinsic matching of the current sources. In the second and third scenarios, the DAC is calibrated at $40^\circ C$ using the conventional technique in [24] and the proposed $g_m$ technique, respectively. After the calibration is done, the temperature is swept and the temperature drift of the DAC performances are obtained.

6.3.1 Current Sources Temperature Drift

The 15 MSB CCSs versus the temperature of three simulation scenarios and the distributions of the temperature drift are shown in Figure 6.4. The temperature drift of a CCS is defined as the difference between $-40^\circ C$ and $120^\circ C$. The calibration technique in [24] forces the 15 MSB CCSs within the CAL DAC resolution at the calibration temperature $40^\circ C$. The currents diverge when the temperature drifts high or low. The $g_m$ calibration technique reduces the distribution sigma to 0.22 LSB across the whole temperature range.

The calibration technique in [24] does not improve the drift compared to the scenario without any calibration because the CAL DAC output current is temperature independent. The $g_m$ calibration technique improves the standard deviation of the temperature drift by a factor of 2.3. The remaining drift is caused by the ignored $(\Delta \beta/\beta)I$ component in the current mismatch.

6.3.2 Linearity

Figure 6.5 shows the INL and DNL of the 14-bit DAC in three scenarios from $-40^\circ C$ to $120^\circ C$. Table 6.1 summarized the worst INL and DNL in the whole temperature range and the INL and DNL at the calibration temperature ($40^\circ C$). Without any calibration, the raw linearity is about 10-bit from the INL perspective. The INL and DNL are slightly better at high temperature due to smaller $g_m$ as shown in Figure 5.2. With the calibration technique in [24], the DACs linearity is better and achieves 14-bit accuracy at $40^\circ C$. But the DAC INL is seven times worse when the temperature changes to $-40^\circ C$, where the DAC INL drops to 12.5-bit. With the proposed $g_m$
calibration technique, the INL and DNL are always within ±0.5 LSB across the whole temperature range and the DAC achieves real 14-bit level accuracy. The $g_m$ calibration technique reduces the
Figure 6.5: Simulation results of INL and DNL v.s. temperature. (a) no calibration. (b) calibration technique in [Mercer JSSC 2007]. (c) $g_m$ calibration technique.

Table 6.1: Summary of INL and DNL (−40°C-120°C)

<table>
<thead>
<tr>
<th></th>
<th>INL (14-bit level LSB)</th>
<th>DNL (14-bit level LSB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Worst case at CAL T</td>
<td>Worst case at CAL T</td>
</tr>
<tr>
<td>No CAL</td>
<td>7.21</td>
<td>6.02</td>
</tr>
<tr>
<td>Mercer JSSC 2007 [24]</td>
<td>1.35</td>
<td>0.20</td>
</tr>
<tr>
<td>$g_m$ CAL</td>
<td>0.46</td>
<td>0.23</td>
</tr>
</tbody>
</table>

worst case INL and DNL of the DAC in the whole temperature range by a factor of 15.7 and 12.8 compared with the intrinsic matching and by a factor of 2.9 and 2.8 compared with the conventional calibration method.

From the area perspective, 1-bit linearity improvement requires quadruple the device area. Therefore, for the same INL, the $g_m$ technique reduces the area of the current sources by a factor of 246 compared with no calibration and by a factor of 8.6 compared with the technique in [24].
6.4 Conclusion

In this chapter, a new foreground calibration technique is presented. The CAL DAC capable of tracking temperature variations is introduced. A transistor level 14-bit DAC is implemented in a standard 65nm CMOS processes to validate this technique. The simulation results show a significant linearity improvement across a wide temperature range from $-40 \degree C$ and $120 \degree C$, compared with the intrinsic matching and the conventional calibration method [24].
Chapter 7

Two-Parameter DAC Calibration Technique

7.1 Introduction

In this chapter, a foreground two-parameter calibration technique is presented to further improve the temperature stability comparing with the $g_m$ calibration technique proposed in Chapter 6. Two different CAL DACs have the same temperature dependence as the two components of the current mismatch $\Delta I$. Therefore, the sum current of the two CAL DACs automatically tracks the temperature variations of the current source mismatch with no need of temperature information. The calibration procedure to obtain the CAL DAC input codes utilizes two bias current settings instead of two temperatures, which reduces the calibration time and expense significantly.

The remainder of this chapter is organized as follows. Section 7.2 proposes the two-parameter calibration technique and the circuit implementation, followed by the silicon measured results in Section 7.3.

7.2 Two-parameter Calibration Technique

7.2.1 Basic Idea

The $g_m$ calibration technique is based on the assumption that $g_m(-\Delta V_{TH})$ dominates $\Delta I$ and $(\Delta \beta/\beta)I$ can be omitted. The FSC of the CAL DAC is proportional to $g_mV_{ref}$, where $V_{ref}$ is a
constant voltage derived from the bandgap voltage. $I_{cor}$ has the same temperature dependence as $g_m(-\Delta V_{TH})$, hence tracks $-\Delta I(T)$ better.

However, it is common practice to reduce $g_m$ in the current source design. The simulated ratio of the standard deviation of $g_m(-\Delta V_{TH})$ to that of $(\Delta \beta/\beta)I$ is 2.4-to-1 for the current sources plotted in Figure 5.3. It can be seen that several of the current sources exhibit a high offset and low slope, indicating $(\Delta \beta/\beta)I$ larger than the value that is usually assumed. Therefore, $(\Delta \beta/\beta)I$ should be taken into account in improving the temperature stability.

The proposed CCS includes two CAL DACs as shown in Figure 7.1 and their sum current produces $I_{cor}$. Since the FSCs of the CAL DACs I and g are proportional to $I$ and $g_m V_{ref}$, they have the same temperature dependence as $(\Delta \beta/\beta)I$ and $g_m(-\Delta V_{TH})$, respectively. The calibration routine will compensate both mismatch components using their corresponding CAL DACs. Once the calibration is completed, $I_{cor}$ will be equal to $-\Delta I(T)$ at any temperature.

At least two data points are needed to determine the two unknown CAL DAC input codes. One option is using two different temperatures, which can be time consuming and expensive for production test. Instead, the proposed calibration procedure uses two bias current settings based on the observation that $g_m$ is a function of $I$ as well. Calibration can, therefore, be done without the need to vary temperature, achieving significant calibration time and cost saving.

### 7.2.2 Calibration Procedure

The proposed iterative procedure includes an initial cycle and $N$ successive cycles. Each cycle except the initial one includes four steps. Figure 7.2 provides a numeric example. $C^I_n$ and $C^g_n$ are the currents of the CAL DACs I and g after $n$-th cycle and their targets are $-(\Delta \beta/\beta)I$ and $g_m V_{ref}$.
CHAPTER 7. TWO-PARAMETER DAC CALIBRATION TECHNIQUE

Figure 7.2: An example of the proposed two-parameter calibration procedure.

\[-\left(g_m(-\Delta V_{TH})\right), \text{ respectively.}\]

The calibration starts with the bias current setting \( I \). During the initial cycle \((n = 0)\), \( C^0_g \) is set to be 0 and only CAL DAC I compensates \( \Delta I \) by measuring against \( I_{ref} \) through the SAR algorithm [24]. After the initial cycle, \( C_0^I \) is equal to \(-((\Delta \beta / \beta)I + g_m(-\Delta V_{TH}))\).

In Step 1 of the first cycle, the bias current is reduced to \( \alpha I \). \( \alpha \) is required to be less than 1 as will be explained later. Two mismatch components scale differently. The current mismatch \( \Delta I \) becomes \( \alpha_I(\Delta \beta / \beta)I + \alpha_g g_m(-\Delta V_{TH}) \), where \( \alpha_I = \alpha \) and \( \alpha_g = \sqrt{\alpha} \). The FSCs of the CAL DACs I and g scale to \( \alpha_I \) and \( \alpha_g \) as well. Then in Step 2, only CAL DAC g is changed to make \( I_{cor} \) compensate \( \Delta I \). In Step 3 the bias current is restored to \( I \) and only CAL DAC I is changed to compensate \( \Delta I \) in Step 4. These four steps repeat \( N - 1 \) cycles.

After Step 2 and 4 of the \( n \)-th cycle, \( C_I^n \) and \( C_g^n \) can be expressed by

\[ \alpha_I C_I^{n-1} + \alpha_g C_g^n = -\left(\alpha_I \frac{\Delta \beta}{\beta} I + \alpha_g g_m(-\Delta V_{TH})\right), \quad (7.1) \]

\[ C_I^n + C_g^n = -\left(\frac{\Delta \beta}{\beta} I + g_m(-\Delta V_{TH})\right). \quad (7.2) \]
Combine (7.1) and (7.2) into the matrix format to obtain
\[
\begin{bmatrix}
1 & 1 \\
0 & \alpha_g
\end{bmatrix}
c_n = \begin{bmatrix}
0 & 0 \\
-\alpha_I & 0
\end{bmatrix} c_{n-1} + \begin{bmatrix}
1 & 1 \\
\alpha_I & \alpha_g
\end{bmatrix} m, \tag{7.3}
\]
where
\[
c_n = \begin{bmatrix}
C_I^n & C_g^n
\end{bmatrix}^T, \tag{7.4}
\]
\[
m = -\begin{bmatrix}
(\Delta \beta / \beta) I & g_m(-\Delta V_{TH})
\end{bmatrix}^T. \tag{7.5}
\]
(7.3) can be written as
\[
c_n = A c_{n-1} + B m, \tag{7.6}
\]
where
\[
A = \begin{bmatrix}
1 & 1 \\
0 & \alpha_g
\end{bmatrix}^{-1} \begin{bmatrix}
0 & 0 \\
-\alpha_I & 0
\end{bmatrix} = \begin{bmatrix}
\alpha_I / \alpha_g & 0 \\
-\alpha_I / \alpha_g & 0
\end{bmatrix}, \tag{7.7}
\]
\[
B = \begin{bmatrix}
1 & 1 \\
0 & \alpha_g
\end{bmatrix}^{-1} \begin{bmatrix}
1 & 1 \\
\alpha_I & \alpha_g
\end{bmatrix} = \begin{bmatrix}
1 - \alpha_I / \alpha_g & 0 \\
\alpha_I / \alpha_g & 1
\end{bmatrix}. \tag{7.8}
\]
Solving (7.6) recursively leads to
\[
c_n = A^n c_0 + \left( \sum_{k=0}^{n-1} A^k \right) B m = \begin{bmatrix}
(\alpha_I / \alpha_g)^n & 0 \\
-(\alpha_I / \alpha_g)^n & 0
\end{bmatrix} c_0 + \begin{bmatrix}
1 - (\alpha_I / \alpha_g)^n & 0 \\
(\alpha_I / \alpha_g)^n & 1
\end{bmatrix} m. \tag{7.9}
\]
If \(\alpha_I / \alpha_g < 1\),
\[
\lim_{n \to \infty} c_n = m, \tag{7.10}
\]
indicating that the CAL DACs I and g converge linearly to their targets. Actually, the calibration procedure forces \(I_{cor}\) equal to \(-\Delta I\) under both bias current conditions. The only way to satisfy this requirement is for each component of the current mismatch to be compensated by its corresponding CAL DAC.

The rate of convergence is \(\alpha_I / \alpha_g\), equal to \(\sqrt{\alpha}\). The error between \(c_n\) and \(m\) keeps shrinking through cycles until it drops below the quantization error of the CAL DAC, i.e. \(\pm 0.5\) LSB. Therefore, the maximum number of cycles for a \(M\)-bit CAL DAC is given by
\[
N = \frac{M}{\log_2(\alpha_I / \alpha_g)}. \tag{7.11}
\]
7.2.3 Calibration Implementation

A 16-bit DAC is designed to test the proposed calibration technique. The first four MSB are thermometer coded, composed of 15 identical CCSs with a nominal current of 2.5 mA each.

The CAL DACs I and g are identical circuits, as shown in Figure 7.3, with their FSCs determined by $I_{in}$. The digital input code controls how much current is steered to the output. Each CAL DAC is 8-bit and its resolution is equal to 0.5 DAC LSB. $I_{in}$ of a CAL DAC I is simply a copy of the main current source $I$, whereas $I_{in}$ of a CAL DAC g is equal to $g_m V_{ref}$ that can be generated by the “$g_m V_{ref}$ current generator” in Section 6.2.3 [10, 13].

The calibration loop configuration is shown in Figure 7.4. During calibration, the cascode switch $M_{CAS1}$ is off and $M_{CAS2}$ is on, thereby steering the current of the CCS to the calibration circuitry that includes a comparator and a calibration control circuit. All the MSB CCSs are calibrated against the on-chip reference current $I_{ref}$. The $V_{TH}$ mismatch between $M_{CAS1}$ and $M_{CAS2}$ causes an additional CCS mismatch through channel-length modulation after the CCS finishes calibration and is switched to normal operation. This additional mismatch can not be compensated by the proposed calibration technique, so it has to be taken care of by design. The simulated standard deviation of the induced CCS mismatch by this mechanism is 0.2 LSB.
The ratio $\alpha$ between two bias current conditions does not need to be accurate and the only requirement for convergence is less than 1. Smaller $\alpha_I/\alpha_g$ indicates faster convergence; however, the transistor $M_{cs}$ will be out of strong inversion region if $\alpha$ is too small. Therefore, $\alpha = 0.25 (\alpha_I/\alpha_g = 0.5)$ is chosen in this implementation and the calibration procedure takes at most 8 cycles according to (7.11).

7.3 Experimental Results

A test chip has been fabricated in a 65 nm standard CMOS process. It can be configured in four different ways: a DAC without calibration, the conventional calibration [24], the $g_{m}$ calibration, and the proposed scheme. Turning off both CAL DACs exhibits the intrinsic matching of the current sources without calibration. The conventional calibration in [24] and the $g_{m}$ calibration are equivalent to only utilizing either CAL DAC I or $g$. The proposed calibration technique uses both CAL DACs.
Any calibration is run only once at 25°C. Then, all the CAL DAC input codes are fixed and the 15 MSB CCSs are measured with temperature sweeping from −40°C to 85°C.

### 7.3.1 MSB CCSs vs. Temperature

Figure 7.5 shows the measured currents of the 15 MSB CCSs and the worst standard deviation across the temperature range for a typical chip. The unit is 16-bit level DAC LSB. The
CHAPTER 7. TWO-PARAMETER DAC CALIBRATION TECHNIQUE

mismatch range without any calibration is 16 LSB at $-40\,^{\circ}\mathrm{C}$, indicating 12-bit level intrinsic matching. The mismatch decreases with increasing temperature, which agrees with the analysis in Section 5.1.

All the three techniques calibrate the MSB CCSs within $4096 \pm 0.6$ LSB at the calibration condition $25\,^{\circ}\mathrm{C}$. The residue mismatches ($\pm 0.6$ LSB) after calibration is caused by three mechanisms. The first is the $V_{TH}$ mismatch of the cascode switches that causes the CCS mismatch through channel-length modulation ($1\sigma = 0.2$ LSB). The other two mechanisms are the CAL DAC resolution ($\pm 0.25$ LSB) and the measurement noise during calibration. With the calibration technique in [24] applied, the currents drift significantly, especially at cold. The currents drift less when using the technique in [10]. The proposed two-parameter calibration technique further reduces the temperature sensitivity, and the MSB CCSs are within $4096 \pm 0.7$ LSB across the whole temperature range.

When comparing the calibration technique in [10] and the proposed one, the worst standard deviation of the MSB CCSs reduces from 0.64 LSB to 0.44 LSB. If the aforementioned three residue mismatches are subtracted, the improvement from the proposed calibration technique should be more significant.

To achieve this level of performance, the current source array would need to be 106 times larger if no calibration were included. If only CAL DAC I were used, as in [24], the current source array would need to be 7.1 times larger. Similarly if only CAL DAC g were used [10], the array would need to be 2.1 times larger. The parasitic capacitance with the proposed calibration technique is reduced significantly, which enables the DAC to achieve superior dynamic linearity at high frequency [13].

The temperature drift of CCS is defined as the maximum current change across the whole temperature range. The standard deviations of the temperature drift for the MSB CCSs are shown in Table 7.1. The proposed two-parameter calibration technique offers improvement compared with the previously published techniques. The remaining temperature drift may arise from the temperature drift of the residue mismatches and possible temperature dependence of $\Delta V_{TH}$ and $\Delta \beta/\beta$.

7.3.2 Calibration Convergence

Figure 7.6 shows all the CAL DAC input codes after each cycle during the two-parameter calibration procedure. The bias current ratio $\alpha$ is 0.25 and $\alpha_I/\alpha_g$ is about 0.5. The codes settle after the 6-th cycle. The final code range of the CAL DAC g is about twice that of the CAL DAC I, which validates the simulated ratio of $g_m(-\Delta V_{TH})$ to $(\Delta \beta/\beta)I$. 

80
Chapter 7. Two-Parameter DAC Calibration Technique

Table 7.1: Measured temperature drift of 15 MSB CCSs

<table>
<thead>
<tr>
<th></th>
<th>Temperature drift 1σ (LSB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No CAL</td>
<td>1.6</td>
</tr>
<tr>
<td>Mercer JSSC 2007</td>
<td>1.6</td>
</tr>
<tr>
<td>g_m CAL</td>
<td>0.8</td>
</tr>
<tr>
<td>Two-parameter CAL</td>
<td>0.6</td>
</tr>
</tbody>
</table>

7.4 Conclusion

The proposed foreground calibration technique resolves the temperature sensitivity issue of the conventional techniques. The two CAL DACs automatically track the temperature variation of the current source mismatch. The measurement results of a 16-bit DAC test chip with the proposed calibration technique show that the MSB currents maintain 16-bit level matching across the temperature range from −40 °C to 85 °C with 12-bit level intrinsic matching. In addition, a test friendly calibration procedure is proposed to make production test cost-effective. The CAL DAC input codes can be stored in a non-volatile memory, eliminating the need for background calibration or re-calibration when the temperature varies.
Figure 7.6: Measured convergence of CAL DAC input codes of 15 MSB CCSs. (a) CAL DAC I. (b) CAL DAC g.
Chapter 8

Conclusion

Several design techniques are proposed to improve the noise and linearity of data converters in this thesis. These techniques break the performance trade-offs by providing powerful new degrees of freedom in design. Therefore, the performance improvements arises from the fundamental design innovations instead of the advancement of the CMOS processes. All the designed techniques were implemented in test chips and validated by the measurements.

The sampling noise reduction technique allows the size of the input capacitor to be determined almost independently of the noise requirement. The technique uses active circuits and a second capacitor not driven by the input to cancel the noise sampled on the input capacitor. Test chip measurements demonstrate that the effective sampled thermal noise can be reduced by as much as 67% without change to the input capacitor.

Traditionally, switched-capacitor amplifiers face tradeoffs among power consumption, noise, and accuracy. The bandwidth-switching technique splits the amplification phase into two sub-phases. In the first sub-phase, the amplifier is configured with high bandwidth and high slew rate. In the second sub-phase, the bandwidth is reduced, achieving the target settling accuracy but with much lower noise. Test chip measurements demonstrate that the proposed THA achieves 45% noise power reduction in the amplification phase as well as improved linearity and lower power consumption.

Foreground calibration techniques improve the matching of current sources of small area, enhancing both static and dynamic linearity. The $g_m$ calibration technique tracks the temperature variations with the new CAL DAC. A transistor level 14-bit DAC is implemented in a standard 65nm CMOS processes to verify this technique. Simulation results show a significant linearity improvement across a wide temperature range from $-40^\circ C$ and $120^\circ C$, compared with the intrinsic
matching and the conventional calibration method [24].

The two-parameter foreground calibration technique further improves the temperature stability. The two CAL DACs automatically track the temperature variation of the current source mismatch. The measurement results of a 16-bit DAC test chip with the proposed calibration technique show that the MSB currents maintain 16-bit level matching across the temperature range from $-40\,^\circ C$ to $85\,^\circ C$ with 12-bit level intrinsic matching. In addition, a test friendly calibration procedure is proposed to make production test cost-effective. The CAL DAC input codes can be stored in a non-volatile memory, eliminating the need for background calibration or re-calibration when the temperature varies.

All the proposed design techniques improve the performance of basic building blocks. Therefore, they are not limited in the data converter design and can be applied to other circuits and systems using the building blocks as well.

In the future research, there is still plenty of design space to explore. The data converters are moving to RF frequency (GHz) steadily. At such high frequency, the impact of the clock jitter noise becomes more and more severe and the thermal noise may not dominate. Research efforts are needed to improve the noise at high frequency. In the meantime, the dynamic errors such as the timing mismatch instead of the static current mismatch start to dominate the DACs’ linearity. Calibration techniques to calibrate the dynamic errors are of interest to designers.
Bibliography


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


