## IOWA State University

# Sampled charge reuse for power reduction in switched capacitor data converters 

Saqib Qayyum Malik<br>Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Electrical and Electronics Commons

## Recommended Citation

Malik, Saqib Qayyum, "Sampled charge reuse for power reduction in switched capacitor data converters " (2006). Retrospective Theses and Dissertations. 1284.
https://lib.dr.iastate.edu/rtd/1284

# Sampled charge reuse for power reduction in switched capacitor data converters 

## by

## Saqib Qayyum Malik

# A dissertation submitted to the graduate faculty in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY 

Major: Electrical Engineering (Microelectronics)

Program of Study Committee:
Randall L. Geiger, Major Professor
Degang J. Chen
Chris C-N Chu
Thomas W. Meyer
Robert J. Weber

Iowa State University
Ames, Iowa
2006

## INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

## UMI

UMI Microform 3217298
Copyright 2006 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346

Ann Arbor, MI 48106-1346

Graduate College Iowa State University

This is to certify that the doctoral dissertation of Saqib Qayyum Malik has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.
Major Professor
Signature was redacted for privacy.
Frofte Major Program

## DEDICATION

To my late father,
Abdul Qayyum Malik

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS ..... vii
CHAPTER 1. GENERAL INTRODUCTION ..... 1
Dissertation organization ..... 1
Capacitor sharing and scaling technique for reduced power in pipelined ADCs ..... 2
Flash ADC architecture ..... 2
Pipeline ADC architecture ..... 3
Time-interleaved ADCs ..... 11
A capacitor sharing technique for RSD Cyclic ADC ..... 13
References ..... 14
CHAPTER 2. CAPACITOR SHARING AND SCALING TECHNIQUE FOR REDUCED POWER IN PIPELINED ADCS ..... 18
Abstract ..... 18
Introduction ..... 18
Operation of a typical pipeline stage ..... 20
Proposed technique ..... 22
Operation ..... 22
Advantages ..... 24
Simulation results ..... 25
Conclusions ..... 25
Acknowledgment. ..... 26
References ..... 26
CHAPTER 3. A CAPACITOR SHARING TECHNIQUE FOR RSD CYCLIC ADC ..... 27
Abstract ..... 27
Introduction ..... 27
Conventional cyclic structure ..... 28
Proposed cyclic structure ..... 31
A. Benefits of the proposed structure ..... 33
B. Design Issues ..... 35
Implementation and simulation results ..... 36
Conclusion ..... 36
References ..... 37
CHAPTER 4. A LOW TEMPERATURE SENSITIVITY SWITCHED-CAPACITOR CURRENT REFERENCE ..... 38
Abstract ..... 38
Introduction ..... 38
Background ..... 39
Current reference architecture ..... 39
Design considerations ..... 41
Increasing the output resistance ..... 41
Hold capacitor and ripple ..... 41
Improving the settling time ..... 43
Simulation results ..... 43
Conclusions ..... 45
Acknowledgements ..... 45
References ..... 45
CHAPTER 5. AREA EFFICIENT LAYOUT STRATEGIES FOR EXTREME-RATIO MOS TRANSISTORS ..... 46
Abstract ..... 46
Introduction ..... 46
Layout comparison method ..... 49
Layout structures ..... 51
Alternating Bar structure ..... 52
Waffle structure ..... 53
Zipper structure ..... 54
Star Zag ..... 56
Fingered-Waffle ..... 57
Hexagonal ..... 58
Performance comparison ..... 59
Complete transistor ..... 61
Alternating Bar Structure ..... 62
Waffle ..... 63
Zipper ..... 65
Star Zag ..... 66
Fingered Waffle ..... 68
Hexagonal ..... 70
Conclusions ..... 70
Acknowledgments ..... 71
References ..... 71
CHAPTER 6. CONCLUSION ..... 72
Capacitor sharing and scaling technique for reduced power in pipelined ADCs ..... 72
Contributions ..... 72
A capacitor sharing technique for RSD cyclic ADC ..... 73
Contributions ..... 73
A low temperature sensitivity switched-capacitor current reference ..... 73
Contributions ..... 74
Area efficient layout strategies for extreme-ratio MOS transistors ..... 74
Contributions ..... 75

## ACKNOWLEDGEMENTS

I am fortunate to have been blessed with family and friends who were always there to help me in times of need. It is impossible to acknowledge everyone individually but I hope the ones I do not mention specifically know of my gratitude.

First, I would like to thank my major professor, Randy Geiger, for guiding me through my graduate studies. Over the years, I have learned a lot from him and owe him my deepest thanks for all his time and patience. It is hard to find a professor like him who gives you so much attention and never hurries you out of his office. I will always remember the many road trips that he organized to conferences all over U.S. The conferences were great academic experiences and the driving on these trips was an adventure in its own. I would also like to thank the rest of my committee, professors Degang Chen, Chris Chu, W. Thomas Meyer, and Robert Weber for agreeing to oversee my work.

Friends and colleagues. Good friends and colleagues help make the graduate school experience bearable. Mark Schlarmann and I had a lot of conversations about a lot of topics. All those discussions helped me better understand topics I thought I already knew all about. Talk about getting an education! Whether publishing papers or just spending time together, it was always good to be in the company of Mezyad Amourah. I had some fun times with him and his adorable daughter, Rama, and great food made by his wife. I also would like to thank my recent office mate Yu Lin for her inspirational hard work and insights. Vipul Katyal has been very helpful in our discussions and my wrapping up of this work. I would also like to thank Basem Soufi, Xin Dai, and Hanqing Xing for providing their ADCs as test bed for the technique in this work. I hope Basem keeps up his energy, enthusiasm, and helpfulness throughout his life. Of the many others, I would like to mention and thank Raza-ul-Mustafa, Ahmed Younis, Mao-Feng Lan, Jing Ye, Yvette Lee, Sudha Nagavarapu, Beatriz Olleta, Yonghui Tang, Yonghua Cong, Ahmed Ismail, Huiting Chen, Jie Yan, Le Jin, Kumar Parthasarathy, K. C. Tiew, Haibo Fei, Su Chao, and Latinus Eddie Boylston for all the memories. The list cannot be complete without mentioning the wonderful lamboo (tall) Ahmed Hashim. He has been like a younger brother and a wonderful friend. I hope to come close in becoming like him in so many aspects of his personality.

My thanks go to all the families I have known in Ames who opened their homes and hearts to me and made me feel at home while I was thousands of miles away from home.

My family, including my mother, my brothers Faisal and Faraz, and my sister Faryal, has been behind me in going to graduate school and finishing it. I am lucky to have Faryal as my wonderful sister. No
praise can do justice to her and to her big heart. Who can ever thank their mother enough? I would not be here and would not have achieved any success without the sacrifices of my mother. I hope to make her happy and keep her happy for as long as I live.

Lastly, I would like to thank my wife, Rabia Moussa, for her companionship and support in finishing my studies. We are thankful to God Almighty for our beautiful daughter, Noor Suraya Malik. I look forward to our life together.

## CHAPTER 1. GENERAL INTRODUCTION

Invention of integrated circuits marked the beginning of the electronic revolution. Electronic devices have become ubiquitous and the trend of miniaturization of transistors has been at the heart of increased functionality available to the consumers. As predicted by Gordon Moore in what is now commonly referred to as the Moore's law [1], the number of transistors on a single chip have roughly doubled every eighteen months. The increased number of transistors on a chip has resulted in more computing power becoming available and has enabled devices such as laptops, digital cameras, PDAs, etc. The digital circuitry benefits tremendously from the constant shrinking of the device sizes; the benefit for analog circuits is not quite so dramatic and, in many instances, the analog design becomes more challenging in newer processes that have lower power supply values. Real world signals such as voice are inherently analog in nature. Consequently, these analog signals are typically converted to their digital equivalent in order to fully utilize the benefits of available digital circuitry. Analog-to-digital converters (ADCs) are used to perform this function.

For mobile battery-powered applications, low power dissipation is a critical requirement. Techniques that can reduce power dissipation or area in ADCs find use in a variety of applications and are of significance to the semiconductor industry. A technique that can reduce power dissipation or area of ADCs based on switched capacitor circuits is presented in this work. The technique is demonstrated to be applicable to pipeline and cyclic ADCs. A chapter describing more efficient use of area for extreme-ratio transistors is also presented.

## Dissertation organization

This dissertation is a collection of four papers that have been published or have been prepared for publication. The first two papers describe a switched capacitor technique to reduce area and power consumption in pipelined and cyclic ADCs. The third paper describes a switched capacitor technique used to obtain a current reference with low sensitivity to temperature variations. The fourth paper describes alternate layout techniques for MOS transistors. Since the first two papers provide limited background, detailed background for ADCs will be presented in this chapter in order to help the reader get a better understanding of the novel idea presented in the papers. The conclusion chapter will summarize the contribution of each work.

## Capacitor sharing and scaling technique for reduced power in pipelined ADCs

ADCs convert an analog signal into its digital equivalent. A 1-bit ADC can be implemented using a comparator, as shown in Figure 1. An analog signal is applied to the input of the comparator. If the input is higher than a reference voltage, $\mathrm{V}_{\text {ref }}$, the comparator output is a ' 1 '; a ' 0 ' is generated otherwise. In typical applications, the number of bits required can be higher. For example, if a digital control signal is needed to be ' 1 ' if the battery in a cell phone is $50 \%$ drained, a 1 -bit ADC would suffice. However, if sensing of the battery status in increments of, say, $5 \%$ were desired, an ADC with more number of bits would be needed.

Table 1 summarizes a few CMOS ADCs appearing in recent years and their respective architectures and salient features. A few commonly used architectures of ADCs are presented next along with their benefits and limitations.


Figure 1. Comparator as a 1-bit ADC and example signals

## Flash ADC architecture

The flash ADC processes the input in a parallel manner to determine the digital representation of the given analog input. A basic 3-bit flash ADC [2] is shown in Figure 2. If the input signal magnitude can vary between 0 and $V_{\text {ref }}, 2^{3}$ resistors are used to generate the required number of equally-spaced reference voltages. The input signal is applied to $2^{3}-1$-bit ADCs (simple comparators), each comparing the input signal with a reference voltage. The output of these comparators is then converted to the desired 3 bits using an encoder.

Table 1. A summary of a few ADCs appearing in the literature

| Bits | MS/s | Process | P(mW) | Architecture | Year | Ref. |
| ---: | ---: | :--- | :--- | :--- | :--- | :--- |
| 12 | 110 | CMOS, $0.18 \mu$ | 97 | Pipeline | 2005 | $[3]$ |
| 14 | 12 | CMOS, $0.18 \mu$ | 98 | Pipeline | 2004 | $[4]$ |
| 10 | 100 | CMOS, $0.18 \mu$ | 67 | Pipeline | 2004 | $[5]$ |
| 13 | 16 | CMOS, $0.25 \mu$ | 78 | Pipeline | 2004 | $[6]$ |
| 8 | 20000 | CMOS, $0.18 \mu$ | 9000 | 80 parallel pipelines | 2003 | $[7]$ |
| 12 | 75 | CMOS, $0.35 \mu$ | 290 | Pipeline | 2003 | $[8]$ |
| 10 | 30 | CMOS, $0.30 \mu$ | 16 | Pipeline | 2003 | $[9]$ |
| 8 | 4000 | CMOS, $0.35 \mu$ | 4600 | 32 parallel pipelines | 2002 | $[10]$ |
| 10 | 120 | CMOS, $0.35 \mu$ | 234 | 2 parallel pipelines | 2002 | $[11]$ |
| 6 | 1300 | CMOS, $0.35 \mu$ | 500 | Flash | 2001 | $[12]$ |
| 6 | 1100 | CMOS, $0.35 \mu$ | 300 | Flash | 2001 | $[13]$ |
| 10 | 20 | CMOS, $0.5 \mu$ | 75 | Subranging | 1999 | $[14]$ |
| 8 | 75 | CMOS, $0.5 \mu$ | 70 | Parallel Pipeline | 1998 | $[15]$ |
| 6 | 200 | CMOS, $0.5 \mu$ | 150 | Folding and Interpolating | 1998 | $[16]$ |
| 10 | 100 | CMOS, $1 \mu$ | 1100 | Parallel Pipeline | 1997 | $[17]$ |
| 12 | 4 | CMOS, $0.8 \mu$ | 45 | Pipeline | 1996 | $[18]$ |
| 13 | 5 | CMOS, $1.2 \mu$ | 166 | Pipeline, 2 bps | 1996 | $[19]$ |
| 8 | 70 | CMOS, 0.8 u | 110 | Folding and Interpolating | 1995 | $[20]$ |
| 10 | 20 | CMOS, $1.2 \mu$ | 20 | Pipeline, 1.5 bps | 1995 | $[21]$ |

Flash ADCs need one clock cycle to convert the analog signal. As a result, they can operate at very high speeds. However, this speed comes at the cost of more hardware resulting in higher power dissipation. To achieve an extra bit of resolution, the hardware needs to be doubled. If the area occupied for one bit is A, then the total area for an $n$-bit flash ADC is approximately $2^{n} \cdot A$. The doubling of hardware roughly corresponds to a doubling of power dissipation. This geometrical increase in area and power dissipation with the increase in number of bits of resolution limits these structures to the 6 to bit range. [12-13].

## Pipeline ADC architecture

As mentioned earlier, the flash ADCs process the analog input in a parallel manner to achieve the conversion in one clock period. Due to this parallel nature, the area and power requirements double with every incremental bit. Instead of processing the input in parallel, a pipeline ADC serializes the conversion process, as shown in Figure 3. The complete pipeline is subdivided into stages with each stage processing the signal from its preceding stage. Each stage can be designed to generate 1 or more bits per stage. The analog signal is applied at the input of the first stage. The stage has a sub-ADC


Figure 2. A 3-bit flash ADC
that determines the digital bits for the stage. The digital bits are then used by a sub Digital-to-Analog Converter (DAC) to add or subtract an appropriately scaled reference voltage from the input. An operational amplifier (opamp) then amplifies the signal and creates a "residue" voltage that is passed on to the next stage. Each stage repeats the process until the residue has been processed by the last stage. The distinguishing feature of the pipeline ADC is the pipelining, the property that it does not have to wait for a conversion to complete before starting a new one. Since each stage processes the signals independently of the following or the preceding stage, the first stage starts to sample the next input after it has passed on its residue to the next stage. The output bits corresponding to a specific $\mathrm{V}_{\text {in }}$ are collected and output correctly by a time alignment block, as shown in Fig. 3a. The collection in correct order can be done by a simple shift register.


Figure 3. Block diagrams of a pipeline $A D C$ and a single stage of the pipelined ADC

For a 1 -bit/stage structure, the gain of the stage, $\mathrm{A}_{\mathrm{k}}$, is ideally 2 and the residue of the stage k can be written as

$$
\begin{equation*}
V_{r e s}=V_{i n . k+1}=A_{k}\left(V_{i n . k}-d_{k} \cdot \frac{V_{r e f}}{2}\right) \tag{1}
\end{equation*}
$$

where $d_{k}$ is the digital bit generated by the sub-ADC. The ideal transfer characteristics of such a stage are shown in figure 4 [22].


Figure 4. Ideal stage transfer characteristics of a 1-bit/stage pipeline ADC

The transfer curve of an actual pipeline stage usually deviates from the ideal one shown above due to the presence of non-idealities in the circuit. These non-idealities, their effects on the performance of the ADCs, and some solutions found in the literature are presented next.

Non-idealities in pipelined ADCs, their effects, and possible solutions
A traditional pipeline stage combines the function of the DAC, amplification, and a sample-and-hold (S/H) for the subsequent stage into one block, commonly referred to as the Multiplying DAC (MDAC). A popular implementation of the MDAC for a nominal 1-bit per stage pipeline ADC stage is shown in Fig. 5 [23]. For a 1-bit/stage design and ignoring parasitic capacitances, the capacitors are nominally equal and the output can then approximately be given by

$$
\begin{equation*}
V_{o u t, k} \approx\left(1+\frac{C_{2}}{C_{1}}\right) \cdot V_{\text {in }}-\frac{C_{2}}{C_{1}} \cdot d_{k} \cdot V_{r e f} \tag{2}
\end{equation*}
$$

If the opamp and the sub-DAC are assumed to be linear, the sources of errors are charge-injection based offset, comparator offset, and gain error due to capacitor mismatches [22]. The effects of these mismatches on the transfer characteristics of a pipeline ADC stage are shown in Fig. 6. The dashed box marks the maximum allowable value of the output. In the presence of these non-idealities, the output of a stage could go beyond the maximum allowable range resulting in an input for the subsequent stage that is beyond its resolvable limits. As a result, missing decision levels appear in the overall performance of the ADC. In case of the comparator offsets, the transfer curve may be shifted instead of being centered in the complete input range. This shift results in missing codes in the overall transfer curve.


Fig. 5 Conventional Switched-capacitor implementation of an MDAC


Fig. 6 Effects of non-idealities on the transfer curve of one stage
The non-idealities described above are undesirable for obvious performance reasons. Fortunately, techniques exist to reduce the effects of these non-idealities and a few will be described here. The first technique uses a gain that is less than the nominal gain of 2 and uses additional stages to achieve desired resolution. Such a technique that also involved calibrating the stages for offset errors, gain errors, and finite gain errors was proposed by Karanicolas et al [22]. The reduced gain helps in avoiding the output of a stage going beyond the maximum limit, i.e., avoid over-ranging. The resulting digital code needs to be converted into base 2 digital code.

An alternate technique implements the redundant signed digit (RSD) approach. This technique sets nominal gain of the stage to 2 and uses extra comparators in the sub-ADC to detect the over-range condition. Proposed by Ginetti et al., [24] the scheme uses three comparators per stage and has the transfer characteristics shown in Fig. 7. Also referred to as the 1.5 bits/stage architecture, each stage outputs two bits of which one bit provides the redundancy used for correcting errors. The actual digital output for the ADC is obtained by adding together all the bits from individual stages. The radix $<2$ and the RSD techniques are referred to as digital error correction techniques.

As opposed to the correction approaches that preemptively avoid over-ranging or account for it through redundancy, "calibration" techniques can be used to measure the non-idealities in the transfer curve of the overall ADC and correct them after conversion. Self calibration techniques such as the ones proposed in [22], [25], and [26], among others, have been successfully used to improve the accuracy of the pipelined ADCs.


Figure 7. Ideal transfer characteristics for $1.5 \mathrm{bits} /$ stage ADCs using error correction

Typical implementation of a stage
As mentioned earlier, an MDAC provides the functionality of $\mathrm{S} / \mathrm{H}, \mathrm{DAC}$, and gain in a pipeline stage. A typical implementation of an MDAC with a gain of 2 is shown in Fig. 5. In a pipeline ADC, a stage is usually followed by an identical stage. For such a case the $\mathrm{k}^{\text {th }}$ stage is shown in sampling phase in Fig. 8(a). Since stage $k+1$ is in the amplify phase, its input capacitor network does not interact with stage k . At the end of sampling phase, the $\mathrm{k}^{\text {th }}$ stage enters the amplify phase resulting in charge transfer from $\mathrm{C}_{2 . k}$ to $\mathrm{C}_{1 . \mathrm{k}}$. The circuit for that phase is shown in Fig. 8(b).

(a) Stage k sampling, $\mathrm{k}+1$ amplifying

(b) Stage k amplifying, $\mathrm{k}+1$ sampling

Figure 8. Sampling networks of two consecutive MDACs in different stages of operation

While stage k is amplifying the input signal and creating a residue voltage at the output, stage $\mathrm{k}+1$ samples that residue signal. During this phase, ignoring the parasitic capacitances of the opamp, the load at the output of the opamp is given by [28]

$$
\begin{equation*}
C_{L ., p p a m p}=\frac{C_{1 . k} \cdot C_{2 . k}}{C_{1 . k}+C_{2 . k}}+\left(C_{1 . k+1}+C_{2 . k+1}\right) \tag{3}
\end{equation*}
$$

where $\mathrm{C}_{\text {L.opamp }}$ is the total load at the output of the opamp. It is worth noting that the capacitors from stage k are in series [28] and, consequently, appear as an equivalent capacitor smaller in size than the smaller of either of $\mathrm{C}_{1 . k}$ or $\mathrm{C}_{2, k}$. On the contrary, the terms in parentheses depicting the sampling capacitors from stage $\mathrm{k}+1$ are in parallel and their sum contributes directly to the total load driven by the opamp. Power reduction techniques that exploit this dependence have been proposed [19] that scale down the capacitors from stage k to $\mathrm{k}+1$. It was shown in [19] that the optimum scaling factor is approximately the interstage gain. For a 1 -bit per stage, the interstage gain is 2 resulting in the nominal capacitor sizes getting halved every stage. For such a stage, if the capacitors $\mathrm{C}_{1}$ and $\mathrm{C}_{2}$ of all stages are nominally equal, (3) reduces to give the load denoted by $\mathrm{C}_{\text {Lopamp.A }}$ as

$$
\begin{equation*}
C_{L . \text { ppamp } . A}=\frac{C}{2}+(2 C)=2.5 C \tag{4}
\end{equation*}
$$

If the capacitor in later stages are scaled by the interstage gain, i.e., 2 , then (3) can be simplified to give the new capacitive load, $\mathrm{C}_{\text {L.opamp.B }}$ as

$$
\begin{equation*}
C_{\text {L.opanp. } B}=\frac{C}{2}+(C)=1.5 C \tag{5}
\end{equation*}
$$

The technique presented in chapter 4 presents a technique that eliminates the contribution of $\mathrm{C}_{1 . k+1}$ and $C_{2, k+1}$ in (3). The technique reuses the charge stored on $C_{1 k}$ and also automatically scales of capacitors in a pipeline ADC by 2 , as proposed in [19]. This reuse of charge through the sharing of capacitors between stages helps achieve even lower power dissipation. Consequently, the load seen at the output of the first opamp in the proposed technique can be written as

$$
\begin{equation*}
C_{\text {L.opamp.prop }}=\frac{C_{1 . k} \cdot C_{2 . k}}{C_{1 . k}+C_{2 . k}} \tag{6}
\end{equation*}
$$

If equal sized capacitors are used in stage $k$, the capacitive load seen by the opamp is given by

$$
\begin{equation*}
C_{\text {L.opamp.prop }}=0.5 \mathrm{C} \tag{7}
\end{equation*}
$$

Since the opamp in stage k in the proposed technique does not have any interaction with the input sampling network of stage $\mathrm{k}+1$ the loading on the opamp remains the same. Assuming the loading due to comparators can be neglected, the reduction in capacitive load of the first opamp can then be found. The total loading seen in first two stages for the conventional and proposed techniques can provide insight in the potential power savings. Using (3), the total capacitive loading in two consecutive stages can be found and is summarized in Table 2 along with the load reduction with conventional approaches.

Table 2. Capacitive load comparison for conventional pipeline and the proposed technique

| Architecture | Capacitor Scaling | Capacitive load | Proposed technique's load reduction |
| :--- | :--- | :--- | :---: |
| Conventional | No scaling | 5 C | $45 \%$ |
|  | By 2 | 2.25 C | $44.44 \%$ |
| Proposed | Inherently by 2 | 2.75 C | N/A |
|  | By 2 | 1.25 C | N/A |

## Effect of noise in a typical gain stage

Noise is a real-world phenomenon that can be the dominant factor for the maximum achievable performance of integrated circuits, including ADCs. The dominant noise sources are the thermal noise and the quantization noise. The thermal noise due to each capacitor C is given by $\mathrm{kT} / \mathrm{C}$ where k is the Boltzmann's constant and T is the absolute temperature. For an m-bit/stage variant of the MDAC of Fig.5, the output noise power is given by [27]

$$
\begin{equation*}
V_{o}^{2}=2^{m} \cdot \frac{k T}{C} \tag{8}
\end{equation*}
$$

The effective number of bits (ENOB) for an ADC can be expressed in terms of its measured signal-to-noise ratio (SNR) as [29]

$$
\begin{equation*}
E N O B=\frac{S N R_{\text {measured }}-1.76}{6.02} \tag{9}
\end{equation*}
$$

The ENOB in (9) calculation assumes a full-scale sinusoidal input and the SNR is the ratio of the power in the fundamental tone of the output and the noise power. This SNR measurement excludes the power in the harmonics of the fundamental. A more stringent definition of ENOB uses signal-to-noise-and-distortion ratio (SNDR) instead of SNR by including harmonics' power in noise.

$$
\begin{equation*}
E N O B(S N D R)=\frac{S N D R_{\text {measured }}-1.76}{6.02} \tag{10}
\end{equation*}
$$

## Time-interleaved ADCs

As the operating speed of flash and pipeline ADCs is increased, the power dissipation also increases. The higher power dissipation limits how fast the conversion can be performed. To enable higher performance than achievable by a single pipeline, Black and Hodges [30] proposed operating multiple ADCs in parallel. The concept of "time-interleaving" is shown in Figure. 9. The sample-and-hold $(\mathrm{S} / \mathrm{H})$ at the input of the structure samples the input. This sampled input is then passed to one of the parallel $A D C s$ for conversion. For $k$ parallel $A D C s$ and the sampling frequency $f_{s}$, each $A D C$ operates at $f_{s} / k$ instead of the full speed. This reduction in speed of individual ADCs helps to relax the specifications and requirements of the components of each ADC. The multiplexer at the output generates the digital stream by collecting the digital output from each ADC in correct order. An alternate way to look at the technique is to observe that if each individual ADC is capable of operating at a maximum frequency of $f_{s}$, the overall conversion rate can be increased by a factor of $k$. Therefore, to get a higher conversion rate, multiple ADCs could be made to operate in parallel. This technique was utilized by Poulton et al. [31] to achieve an ADC system in GaAs operating at 1GSamples/s. The same technique was utilized by placing successive approximation ADCs in parallel [32] and pipeline ADCs [33] in parallel. Poulton et al. [7] demonstrated an 8bit 20GSamples/s ADC in standard CMOS by time interleaving 80 pipeline stages.


Figure 9. Time interleaved ADC structure

Although an attractive approach to increase the conversion speed of the ADC system, the time interleaved technique does have its issues. In [34] and [35], it was shown that mismatches among the paths result in tones in the frequency response of the system. In the presence of offset mismatchesamong k time-interleaved channels, frequency tones appear at the following frequencies in the output spectrum [36]

$$
\begin{equation*}
\frac{f_{s}}{k} \cdot K, \text { where } K=1,2, \ldots k-1 \tag{11}
\end{equation*}
$$

If the offset of a channel is assumed to be a random variable with normal distribution, zero mean, and variance $\sigma_{o}^{2}$, the resulting SNDR is then given by [36]

$$
\begin{equation*}
S N D R=20 \cdot \log \left(\frac{1}{\sigma_{o}}\right) \tag{12}
\end{equation*}
$$

In case of mismatches among gain $a_{k}$ of different channels, the distortion tone location and SNDR are respectively given by [36]

$$
\begin{gather*}
f_{i n}+\frac{f_{s}}{k} \cdot K, \text { where } K=1,2, \ldots k-1  \tag{13}\\
S N D R=20 \cdot \log \left(\frac{a}{\sigma_{a}}\right)-10 \cdot \log \left(1-\frac{1}{k}\right) \tag{14}
\end{gather*}
$$

where it is assumed that $\mathrm{a}_{k}$ is normally distributed random variable with zero mean and variance $\sigma_{a}^{2}$. The third source of errors in time-interleaved structures is the timing mismatches among the channels. If this timing skew is assumed to be normally distributed random variable with zero mean and variance $\sigma_{t}^{2}$, the output spectrum will have tones present in the response. These tones and the SNDR are respectively given by [36]

$$
\begin{gather*}
f_{\text {in }}+\frac{f_{s}}{k} \cdot K, \text { where } K=1,2, \ldots k-1  \tag{15}\\
S N D R=20 \cdot \log \left(\frac{1}{\sigma_{i} \cdot 2 \pi \cdot f_{\text {in }}}\right)-10 \cdot \log \left(1-\frac{1}{k}\right) \tag{16}
\end{gather*}
$$

For the technique applicable to pipeline ADCs presented in this work, the two capacitor networks used to sample the input could be viewed as the inputs of two time-interleaved ADCs. The mismatch between the two capacitor networks will result in the gain mismatches such as those associated with
the time-interleaved technique mentioned above. These mismatches as well as timing mismatches between two paths will result in tones in the output spectrum degrading the SNDR and ENOB of the ADC. These mismatches resulting in SNDR degradation will be a limiting factor in the maximum achievable resolution for the proposed technique.

## A capacitor sharing technique for RSD Cyclic ADC

As explained earlier, pipelined ADCs can be used to convert an analog signal to digital by serializing the conversion operation. During a complete clock cycle, the residue from each stage is passed down to the next one to finish the conversion. This is achieved by replicating the same stage or modified copy of a stage resulting in more area and power consumption. For applications that require low power dissipations and small footprint, it is possible to take only one stage of an n-stage pipeline ADC and reuse it. If the residue from the stage is fed back to its own input, the ADC could perform the same conversion as the pipeline ADC but would require more cycles to perform one conversion. The ADCs that use this technique are commonly referred to as cyclic, algorithmic, or recycling ADCs.


Figure 10. Block diagram of a cyclic ADC

The first implementation of the cyclic ADC principle was presented by Hornak [37] although the implementation was not completely monolithic. A monolithic cyclic ADC was presented by McCharles et al. [38]. The block diagram of a cyclic ADC is shown in Fig. 10. Traditionally, the S/H function is implemented using a switched capacitor circuit and the remaining function using another switched capacitor based MDAC. The MDAC is typically similar to or identical to the MDACs used in pipeline ADCs . Since the cyclic ADC is similar to a single stage of a pipeline ADC , its performance is also subject to the sources of errors described earlier for the pipeline ADC.

To overcome the comparator offset errors, it is possible to modify the ADC to use the error correction techniques described for the pipeline ADC earlier. Such an approach was presented by Ginetti et al. in [24]. In [39], Garrity and Rakers recognized that by using digital correction and eliminating the need of a dedicated $\mathrm{S} / \mathrm{H}$, the input $\mathrm{S} / \mathrm{H}$ could be converted to a complete stage by adding another set of comparators and sub-DAC. This addition resulted in reducing the total conversion time for the ADC from $n$ to $n / 2$. In a further reuse of hardware, the same authors [40] shared the opamp as well as the comparators. The charge reusing technique described earlier in this chapter for pipeline ADCs can be extended and applied to cyclic ADCs. Such an extension [41] reusing the charge on a feedback network of capacitor in one cycle and reusing it in the next cycle is presented in chapter 3.

## References

[1] Moore, G., "Cramming more components onto integrated circuits," Electronics, April 19, 1965.
[2] R. L. Geiger, P. E. Allen, N. R. Strader, VLSI Design Techniques for Analog and Digital Circuits, McGraw Hill, 1990. ISBN 0-07-023253-9
[3] Andersen, T. J., et al., "A cost effective high-speed 12-bit pipeline ADC in $0.18 \mu \mathrm{~m}$ Digital CMOS," IEEE Journal of Solid-State Circuits, vol. 40, no. 7, pp. 1506-1513, July 2005
[4] Chiu, Y., Gray, P. R., and Nikolic, B., "A 14-b 12-MS/s CMOS pipeline ADC with over 100-dB SFDR," IEEE Journal of Solid-State Circuits, vol. 39, no. 12, pp. 2139-2151, December 2004
[5] Li, J., and Moon, U. -K., "A 1.8-V 67-mW 10-bit 100-MS/s Pipelined ADC Using Time-Shifted CDS Technique," IEEE Journal of Solid-State Circuits, vol. 39, no.9, pp. 1468-1476, September 2004
[6] Liu, M., -H., et al., "A low voltage-power 13-bit 16 MSPS CMOS pipelined ADC," IEEE Journal of Solid-State Circuits, vol.39, no.5, pp. 834-836, May 2004
[7] Poulton, K. et al., "A 20GS/s 8 b ADC with a 1 MB memory in $0.18 \mu \mathrm{~m}$ CMOS," Proceedings of the International Solid-State Circuits Conference, pp. 318-319, 2003
[8] Murmann, B., Boser; B., "A 12b 75MS/s pipelined ADC using open-loop residue amplification," IEEE International Solid-State Circuits Conference, pp. 328-329, February 2003
[9] Miyazaki, D., Kawahito, S., and Furuta, M., "A 10-b 30MS/s low-power pipelined A/D converter using a pseudo-differential architecture," IEEE Journal of Solid-State Circuits, vol. 38, no. 2, pp. 369-373, February 2003
[10] Poulton, K., et al., "A 4GSample/s 8 b ADC in $0.35 \mu \mathrm{~m}$ CMOS", IEEE International Solid-State Circuits Conference, pp. 166-167, February 2002
[11] Jamal, S. M., "A 10-b 120-Msample/s time-interleaved analog-to-digital converter with digital background calibration," IEEE Journal of Solid-State Circuits, vol. 37, no. 12, pp. 1618-1627, December 2002
[12] Choi, M. and Abidi, A. A., "A 6b 1.3Gsamples/s A/D converter in $0.35 \mu \mathrm{~m}$ CMOS," Proceedings of the IEEE International Symposium on Circuits and Systems, 2001
[13] Geelen, G., "A 6b 1.1Gsample/s CMOS A/D Converter," Proceedings of IEEE International Symposium on Circuits and Systems, 2001
[14] Brandt, B., Lutsky, J., "A 75mW 10b 20Msample/s CMOS Subranging ADC with 59dB SNDR," ISSCC Digest of Technical Papers, Vol. 42, 1999
[15] Bright, W., " 8 b 75 Msample/s 70mW Parallel Pipelined ADC Incorporating Double Sampling," ISSCC Digest of Technical Papers, pp. 146-147, Vol. 41, 1998
[16] Jiang, X., Wang, Y., Willson Jr., A., "A 200MHZ 6-Bit Folding and Interpolating ADC In 0.5gm CMOS," IEEE International Symposium on Circuits and Systems, pp. I-5-I-8, 1998
[17] Kim, K.Y., Kusayanagi, N., Abidi, A., "A 10-b, 100-MS/s CMOS A/D Converter," IEEE Journal of Solid-State Circuits, Vol. 32, pp. 302-31, March 1997
[18] Yang, J., Lee, H., "A CMOS 12-bit 4MHz Pipelined A/D Converter with Commutative Feedback Capacitor," IEEE Custom Integrated Circuits Conference, pp. 427-430, 1996
[19] Cline, D. W., Gray, P. R., "A Power Optimized 13-b 5 Msample/s Pipelined Analog-to-Digital Converter in $1.2 \mu \mathrm{~m}$ CMOS," IEEE Journal of Solid-State Circuits, Vol. 31, pp. 294-303, March 1996
[20] Nauta, B., Venes, A., "A 70-MS/S $110-\mathrm{mW}$ 8-b CMOS Folding and Interpolating A/D Converter," IEEE Journal of Solid-State Circuits, Vol. 30, pp. 1302-1308, December 1995
[21] Cho, T. B., Gray, P., "A 10b, $20 \mathrm{Msample} / \mathrm{s}, 35 \mathrm{~mW}$ Pipeline A/D Converter," IEEE Journal of Solid-State Circuits, Vol. 30, pp. 166-172, March 1995
[22] Karanicolas, A., Lee, H. and Bacrania, K., "A 15-b 1-Msample/x Digitally Self-Calibrated Pipeline ADC", IEEE Journal of Solid-State Circuits, vol. 28, pp. 1207-12-5, Dec. 1993
[23] R. J. Baker, H. W. Li, D. E. Boyce, CMOS Circuit Design, Layout, and Simulation, IEEE Press, 1997. ISBN 0-7803-3416-7
[24] Ginetti, B., Jespers, P. G. A., Vandemeulebroecke, A., "A CMOS 13-b Cyclic RSD A/D converter", IEEE Journal of Solid-State Circuits, vol. 27, no. 7, pp. 957-966, July 1992
[25] Lee, H. S., Hodges, D. A., and Gray, P. R., "A Self-Calibrating 15 Bit CMOS A/D Converter," IEEE Journal of Solid-State Circuits, vol. SC-19, pp. 813-819, December 1984
[26] Soenen, E. and Geiger, R. L., "An Architecture and an Algorithm for Fully Digital Correction of Monolithic Pipelined ADC's," IEEE Trans. on Circuits and Systems H, pp. 143-153, March 1995
[27] Azizi, M. Y., et al., "Thermal noise analysis of multi-bit SC gain-stages for low-voltage high resolution pipeline ADC design," Proceedings of the IEE international symposium on signal, circuits, and systems, vol. 2, pp. 365-368, July 2003
[28] C. S. G. Conroy, "A high-speed parallel pipeline A/D converter technique in CMOS," Memorandum no. UCB/ERL M95/94, Electronics Research Laboratory, U. C. Berkeley, November 1994
[29] Baker, R. J., CMOS:Mixed-Signal Circuit Design, volume II, IEEE press, 2002. ISBN 0-471-22754-4
[30] Black, W.C. and Hodges, D.A., "Time interleaved converter arrays,", IEEE J. Solid-State Circuits, Vol. SC-15, pp. 1022-1029, 1980
[31] Poulton, K., Corcoran, J. J., Hornak, T., "A 1-GHz 6-bit ADC system," IEEE J. Solid-State Circuits, Vol. SC-22, no. 6, pp. 962-970, 1987
[32] Yuan, J. and Svensson., C., "A 10-bit 5-MS/s successive approximation ADC cell used in a 70-MS/s ADC array in 1.2- m CMOS," IEEE Journal of Solid-State Circuits, Vol. 29, no. 8, pp. 866-872, August 1994
[33] Conroy, C., Cline, D., Gray, P., "An 8-b 85-MS/S Parallel Pipeline A/D Converter in 1- $\mu \mathrm{m}$ CMOS," IEEE Journal of Solid-State Circuits, Vol. 28, pp. 447-454, April 1993
[34] Jenq, Y. -C., "Digital spectra of nonuniformly sampled signals: Fundamentals and high-speed waveform digitizers," IEEE Transactions on Instrumentation and Measurement, vol. 37, no. 2, pp. 245-251, June 1988.
[35] Petraglia, A. and Mitra, S. K. "Analysis of mismatch effects among A/D converters in a timeinterleaved waveform digitizer," IEEE Transactions on Instrumentation and Measurement, vol. 40, pp. 831-835, Oct. 1991.
[36] Gustavsson, M., Wikner, J. J., and Tan, N. N., CMOS Data Converters for Communications, Kluwer Academic Publishers, 2000. ISBN 0-7923-7780-X
[37] Hornak, T. and Corcoran, J. J., "A High Precision Component-Tolerant A/D Converter," IEEE Journal of Solid-State Circuits, vol. SC-10, No. 6, pp. 386-391, December 1975
[38] McCharles, R., Saletore, V., Black, W., Jr., and Hodges, D., "An algorithmic analog-to-digital converter," Proceedings of the International Solid-State Circuits Conference, pp. 96-97, 1977
[39] Garrity, D. and Rakers, P., "A 10 bit, $2 \mathrm{Ms} / \mathrm{s}$, 15 mW BiCMOS cyclic RSD A/D converter", Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting, pp. 192-195, 1996
[40] Garrity, D. and Rakers, P., "Low power cyclic A/D converter", U.S. Patent 6,535,157, March 2003
[41] Soufi, B., Malik, S. Q., and Geiger, R. L., "A capacitor sharing technique for RSD cyclic ADC," Proceeding of the IEEE Midwest Symposium on Circuits and Systems, August 2005.

# CHAPTER 2. CAPACITOR SHARING AND SCALING TECHNIQUE FOR REDUCED POWER IN PIPELINED ADCS 

A paper published in the Proceedings of the<br>2005 Semiconductor Research Corporation TECHCON

Saqib Q. Malik and Randall L. Geiger


#### Abstract

A technique for reducing power dissipation in pipeline Analog-to-Digital converters (ADCs) is presented. The technique stems from the observation that the amplifier of a given stage is also expected to perform the sample and hold operation by charging the input capacitors of the subsequent stage. At the end of the amplification phase, the feedback capacitor of the first stage holds the residue voltage and it can be directly used as the input signal to the next stage thus eliminating the traditional need for charging an additional input capacitor. This capacitor reuse reduces the total capacitance that a given stage must drive thus reducing the power requirements for the operational amplifiers.


## Introduction

With newer process technologies to fabricate integrated circuits, the number of transistors that fit on a single die has roughly followed the Moore's law. The higher number of transistors has allowed more digital circuits to be realized on a chip resulting in high processing power. On the other hand, analog circuits do not scale as well and the technology scaling benefits have not been quite so dramatic. In order to utilize the available processing power on a chip, it is essential to convert the analog signals to digital. Analog-to-Digital converters (ADCs) form the bridge between the analog and the digital worlds.

ADCs are used in a variety of energy sensitive applications such as mobile devices including mobile phones, personal organizers etc. Due to their dependence on batteries, much effort has been made to minimize the power consumption of these devices. Many techniques have been proposed to reduce the power consumption of the digital as well as the analog part of the systems that make up these mobile devices. The focus of this paper is to present a technique that can reduce the power consumption of the ADCs.

Many architectures for ADCs have been presented over the years. Pipeline ADCs stand out for achieving high speed at high resolution. The block diagram of a pipeline ADC is shown in Fig. 1. The


Figure 1. Block diagram of (a) an n-bit m-bits/stage pipeline ADC (b) a single stage
complete pipeline is subdivided into stages with each stage processing the signal from its preceding stage. Each stage may be designed to generate 1 or more bits per stage. The analog signal is applied at the input of the first stage. The stage has a sub-ADC that determines the digital bits for the stage. The digital bits are then used by a sub-Digital-to-Analog Converter (DAC) to add or subtract an appropriately scaled reference voltage from the input. An operational amplifier (opamp) then amplifies the signal and creates a "residue" voltage that is passed on to the next stage. Each stage repeats the process until the residue has been processed by the last stage. The distinguishing feature of the pipeline ADC is the pipelining, the property that it does not have to wait for a conversion to complete before starting a new one. Since each stage processes the signals independently of the following or the preceding stage, the first stage starts to sample the next input after it has passed on its residue to the next stage. The output bits corresponding to a specific $\mathrm{V}_{\mathrm{in}}$ do need to be aligned in time, as shown in Fig. la. Furthermore, pipeline ADCs lend themselves to calibration algorithms [1]-[3] and correction techniques [4] that allow correcting of many common errors.

A traditional pipeline stage combines the function of the DAC, amplification, and a sample-and-hold (S/H) for the following stage into one block, commonly referred to as the Multiplying DAC (MDAC). By observing that there is interaction between the MDAC of stage $k$ and the input of stage $k+1$, the possibility arises of sharing the capacitors between two adjacent stages in order to save power. A
technique is presented in this paper that exploits the interaction of two stages to reduce the power dissipation in the MDAC of the first stage. Although the sharing can be extended throughout the pipeline, the first two stages generally consume more power than the subsequent stages and hence only those are considered for modification in this work.

Operational amplifiers (opamps) are the major contributors towards the overall power dissipation in pipeline ADCs. In most opamp architectures used in pipelined $A D C s$, the opamp power dissipation is proportional to the capacitive load it must drive. In a typical MDAC based pipeline stage, the opamp must simultaneously drive its own capacitive feedback network as well as the capacitive sampling network of the following stage. In the proposed strategy, a portion of the feedback network is used as the sampling network of the subsequent stage thus eliminating the need to charge a separate sampling network and thus reducing the capacitive drive requirement of the opamps.

The operation of a typical MDAC will be presented in Section II. Section III will describe the details of the proposed technique.

## Operation of a typical pipeline stage

A typical implementation of the MDAC for a nominal 1-bit per stage pipeline ADC stage is shown in Fig. 2 [5]. The two capacitors $C_{1}$ and $C_{2}$ form the sampling network. The two non-overlapping clocks $\phi_{1}$ and $\phi_{2}$ are used to determine the sampling or the amplification mode of the MDAC. For the 1-bit per stage configuration and gain of 2 , nominally $C_{1}=C_{2}$ and $h=0.5$.

For the $\mathrm{k}^{\text {th }}$ stage in the pipeline, the circuit operates as follows. When $\phi_{1}$ goes high, the sampling capacitors sample the input voltage. At the same time, the sub-ADC (not shown) compares the value of the input and generates an output bit $d_{k}$. The digital bit, $\mathrm{d}_{\mathrm{k}}$, is output and sent to the sub-DAC of the stage as well. The output of the DAC, $V_{D A C}$ is $V_{r e f}$ or zero for a $d_{k}$ of 1 or 0 , respectively. At the end of the sample phase, $\phi_{l}$ goes low and $\phi_{2}$ goes high. The resulting configuration is shown in Fig. $3 b$. Charge is transferred from capacitor $C_{2}$ to $C_{I}$ resulting in the residue voltage given by

$$
V_{r e s, k} \approx\left(1+\frac{C_{2}}{C_{1}}\right) \cdot V_{i n, k}-\frac{C_{2}}{C_{1}} \cdot d_{k} \cdot V_{r e f}
$$

While stage $k$ is in the amplification phase, the stage $k+l$ is in sampling phase, as shown in Fig. 3b by the capacitors $C_{l, k+1}$ and $C_{2, k+1}$. By the end of $\phi_{2}$, the opamp output has settled to the residue voltage


Figure 2. A typical switched-capacitor MDAC structure and its clocks


Figure 3. Operation of a typical MDAC (a) Sampling phase (b) Amplifying phase
and the sampling capacitors of stage $\mathrm{k}+1$ have now acquired the residue voltage. A key observation is that at this moment, $C_{l, k}$ also holds the final residue voltage across its terminals.

## Proposed technique

## Operation

The proposed technique exploits the fact that the feedback capacitor, $C_{l, k}$, of the MDAC holds the residue signal at the end of the amplify phase. We propose re-using the charge stored on the feedback capacitor instead of re-sampling the residue on the capacitors of the next stage. To achieve this, $C_{l, k}$ is implemented as a compound capacitor network ( CCN ) of capacitors and switches, as shown in Fig. 4. Each CCN is made up of two capacitors and two switches. For the case where it is desired to have $C_{1}=C_{2}$, each sub-capacitor is made to be equal to $C_{1 . k} / 2$. To use the CCN as a single capacitor, the control clock $\phi$ is set to high placing the two capacitors in parallel. The equivalent capacitance seen between the top terminal, T , and the bottom terminal, B , of the resulting network is given by $C_{1, k}$. When the control clock signal $\phi$ goes low, the sub-capacitors become independently accessible through the terminals FB and DAC. The terminals FB and DAC connect to the next stage's opamp's feedback node and the VDAC respectively.

In the proposed scheme, two CCNs are required for two consecutive stages. Two clocks, chA and chB, in conjunction with other control signals (not shown for simplicity), are used to choose appropriate capacitors $C_{A}$ or $C_{B}$ between two adjacent stages. The circuit uses clocks similar to those used in the traditional case and are shown in Fig. 5d.


Figure 4. Compound Capacitor Network (CCN) and its symbol


Figure 5. Structure of the pipeline with CCNs in different stages of operation

At the beginning of a conversion, the first stage is configured to use the $\mathrm{CCN} C_{A}$, as shown in Fig. 5a. With $c h A$ and $\phi_{1}$ high, $C_{A}$ appears to be a single capacitor and samples the input voltage with $C_{2}$. With $\operatorname{ch} A$ staying high, $\phi_{1}$ goes low and $\phi_{2}$ goes high placing the first stage in amplification mode.

Notice from the configuration of the circuit in this mode, shown in Fig. 5b, that the next stage is not connected to the output of the first stage's opamp at all. This is possible since the input capacitive network of the second stage has been eliminated. As a result, the capacitive load of the opamp is reduced by the total capacitance of the second stage's sampling network.

By the time $\phi_{2}$ becomes low, the appropriate residue voltage has been formed across $\mathrm{C}_{\mathrm{A}}$. At the end of $\phi_{2}, \operatorname{ch} A$ goes low and $\operatorname{ch} B$ goes high. Simultaneously, $\phi_{1}$ goes high placing the first stage in sample mode but using $C_{B}$ instead. This switching of $C_{A}$ into the second stage results in "unfolding" of the CCN . Consequently, when the clock $\phi_{2}$ for the second stage goes high, $C_{A}$ is connected to the second stage's opamp such that it is identical to the topology of the stage in its traditional amplification mode. At the end of $\phi_{2}, \operatorname{ch} A$ goes high again placing $C_{A}$ into the first stage. Similarly, $C_{B}$ is now moved to the second stage and the process is repeated.

## Advantages

The proposed scheme has several benefits when compared to the traditional pipelined ADC. First, since the total capacitance that needs to be driven by the opamp is reduced, the power dissipation requirement of the opamp is reduced. It is well known that the capacitor sizes in a pipelined ADC can be scaled down as one moves towards lower LSB stages for reducing power dissipation [6] and area while still maintaining acceptable overall noise performance. With the sizing strategy described in the proposed structure, the capacitors are scaled by a factor of 2 when going from one stage to the next. This tapering of capacitor sizes provides acceptable noise performance and also provides for a reduction in overall power dissipation. Additionally, since the $\mathrm{k}+1^{\text {th }}$ stage does not need to have its opamp present to sample the residue from stage $k$, opamp sharing between stages $k$ and $k+1$ can be used to further reduce the power requirements. Finally, if the opamp is not modified to reduce its power dissipation, faster speeds of operation become possible since the load seen by the opamp of stage k is reduced.


Figure 6. Simulation results showing the correct synchronized digital outputs for a ramp input

## Simulation results

To validate the proposed technique, a simple setup was selected. The first two stages were implemented using the CCNs followed by two stages of conventional architecture. Each stage's capacitor sizes were designed for a gain of 2 . Behavioral descriptions of circuit elements using Verilog-A were used for the opamps and comparators. The simulation results showing the 4 output bits for a ramp input are shown in Fig. 6. As can be seen, the circuit worked according to design demonstrating that the residues were formed and passed through the pipeline as desired.

## Conclusions

A new technique to reduce the power dissipation in pipeline ADC was presented. In this technique, the feedback capacitor in the first stage is configured to serve as the sampling network of the second stage. This capacitor sharing reduces the capacitive loading on the opamp. Simulation results were
used to prove the soundness of the proposed technique. The technique has also been adapted to work in cyclic ADCs [7].

## Acknowledgment

This work was supported in part by Semiconductor Research Corporation (SRC).

## References

[1] A. Karanicolas and H. Lee, "Digitally Self-Calibrating Pipeline Analog-to-Digital Converter", US Patent No. 5499027, Mar. 1996.
[2] E. Soenen and R. Geiger, "Accuracy Bootstrapping", US Patent No. 5327129, July 1994.
[3] J. Ingino and B. Wooley, "A Continuously Calibrated 12-b, 10-MS/s, 3.3-V A/D Converter", IEEE J. Solid-State Circuits, Vol. SC-33, pp. 1920-1931, Dec. 1998.
[4] S. H. Lewis and P. R. Gray, "A pipelines 5Msamples/s 9-bit analog-to-digital converter," IEEE J. Solid-State Circuits, vol. 22, pp. 954-961, Dec. 1987.
[5] R. J. Baker, CMOS: Mixed-Signal Circuit Design vol. II, IEEE press, 2002.
[6] D. W. Cline, and P. R. Gray, "A power optimized 13-b 5 Msamples/s pipelined analog-to-digital converter in $1.2 \mu \mathrm{~m}$ CMOS", IEEE Journal of Solid-State Circuits, vol. 31, pp. 294-303, March 1996.
[7] B. Soufi, S. Q. Malik, and R. L. Geiger, "A capacitor sharing technique for RSD cyclic ADC", in Proc. of IEEE Midwest Symposium Circ. Systems, August 2005.

# CHAPTER 3. A CAPACITOR SHARING TECHNIQUE FOR RSD CYCLIC ADC 

A paper published in the Proceedings of the<br>2005 Midwest Symposium on Circuits and Systems<br>Basem Soufi, Saqib Q. Malik, and Randall L. Geiger


#### Abstract

A new cyclic ADC structure based on capacitor sharing is presented. This technique reduces the die area of the capacitors in the switched capacitor network by up to $\mathbf{5 0 \%}$. As a result, the proposed scheme also significantly reduces the power consumption requirement of the operational amplifier. This is achieved while maintaining the thermal noise performance and conversion rate of the conventional structure. A $10-\mathrm{bit}, \mathbf{2 . 3 \mathrm { MHz }}$ cyclic ADC using the new structure is implemented in $0.5 \mu \mathrm{~m}$ CMOS. Spectre simulation results of the new structure are presented.


## Introduction

A proliferation of portable devices such as laptop computers, mobile phones, personal digital organizers and digital music players has occurred in recent years. Due to their mobility, portable devices are battery powered. Although the density of digital integrated circuits on a chip has roughly followed Moore's law, battery capacities have not scaled as dramatically. Consequently, power efficient architectures must be used in digital as well as mixed signal circuits such as the analog-to digital converters (ADCs).

ADCs enable the processing of real world analog signals in the digital domain and are ubiquitous in modern portable devices. Of the many ADC architectures, cyclic (or algorithmic) ADCs have the ability to perform analog to digital conversion with minimum area and low power at low to moderate frequencies. Traditional implementations of cyclic ADCs [8] involve a sample-and-hold (S/H) structure along with a gain stage, comparator, and sub-DAC. The hardware is used to implement a form of the binary search algorithm which takes n cycles to produce n -bit digital outputs with a one bit/stage structure. Other architectures combine the Redundant Sign Digit (RSD) technique [9], which compensates for loop offsets, with the inherent S/H functionality in Switched Capacitor (SC) amplifiers to replace the $\mathrm{S} / \mathrm{H}$ in the traditional implementation with another gain stage [10]. Hence, with the addition of another sub-DAC and comparison block, the number of cycles required to
produce n -bit digital output is reduced to $n / 2$ cycles. Usually, each SC gain stage requires a separate operational amplifier (opamp); however, during the sampling phase of the SC amplifier, the opamp is not utilized. By capitalizing on this unused interval of the opamp, a single opamp can be shared between the two stages similar to what is reported in [11]. We will refer to this structure as the 'conventional structure' and discuss it further in Section II.

Two key observations are made on the operation SC amplifier and two-stage cyclic structure. In [12] it is observed that the output voltage of a SC amplifier is actually held across its feedback capacitor. By treating this voltage as a "sampled voltage" for the SC gain stage, the need for an additional sampling capacitor can be eliminated [12]. The observation on the conventional two-stage cyclic structure is during the input sampling phase the second SC network capacitors are not utilized for a useful purpose. By utilizing the second stage capacitors in the input sampling phase, the technique of [12] is implemented in the proposed two-stage cyclic structure without having the drawback of alternatively sampling the input on different capacitors, as is the case in [12]. The proposed architecture automatically provides a limited degree of capacitance scaling similar to the concept of stage scaling of pipeline ADCs discussed in [13]. Hence, the new cyclic ADC is of smaller die area and power consumption levels than the conventional structures. The proposed structure is discussed in Section III. Simulations results are shown in Section IV.

## Conventional cyclic structure

A conventional, two-stage RSD cyclic ADC structure [10] is shown in Fig. 1 where it is assumed that the full-scale input range is $\pm \mathrm{V}_{\text {ref }}$. A typical implementation of the SC network using what is commonly termed a 'flip-around amplifier' is shown in Fig. 2. It consists of two sets of capacitors $\mathrm{CC1}_{\mathrm{a}}, \mathrm{CC1}_{\mathrm{b}}$ and $\mathrm{CC} 2_{\mathrm{a}}, \mathrm{CC}_{\mathrm{b}}$, switches, and a single shared opamp. The conventional cyclic operation is illustrated in Fig. 3 showing the states the SC networks are switched to and their sequence. For every clock phase, the ADC executes the following recursive function:

$$
\begin{equation*}
V_{i+1}=2 \cdot V_{i}-D_{i} \cdot V_{r e f} \tag{1}
\end{equation*}
$$

where the index $i$ denotes the $i^{\text {th }}$ conversion cycle after a sample is taken, $\mathrm{V}_{\mathrm{i}}$ is the $\mathrm{i}^{\text {th }}$ residue voltage seen at the opamp output, $i \in[1, n-1], V_{1}=V_{i n}$, and $D_{i} \in[-1,0,+1]$. The value of $D_{i} \times V_{\text {ref }}$, the DAC output, is a function of the digits $b_{0}$, $b_{1}$ which are the comparison results of $V_{i}$ against two reference values usually set at $\pm V_{\text {ref }} / 4$ as shown in Fig. 1. The value of $D_{i}$ is -1 if neither comparator is set, it is 0 if the lower comparator is set but the upper comparator is unset, and it is +1


Figure 1. Two-stage RSD cyclic ADC structure


Figure 2. Simplified SC networks and clocks for conventional cyclic ADC

(a) Initial State

(b) State A


Figure 3. Switching sequence of conventional SC networks
if both comparators are set. For the switched-capacitor implementation shown in Fig. 2, $\mathrm{V}_{1}$ (the input voltage $\mathrm{V}_{\mathrm{in}}$ ) is sampled on the capacitor pair $\mathrm{Cl}_{\mathrm{a}}, \mathrm{Cl}_{\mathrm{b}}$ at the start of the conversion. In subsequent conversion clock phases, the residue voltage is alternatively sampled on capacitor pairs C2a, C2b and $\mathrm{C} 1 \mathrm{a}, \mathrm{C} 1 \mathrm{~b}$ respectively. The signed codes $\mathrm{b}_{0}, \mathrm{~b}_{1}$ generated at the end of each clock phase are synchronized (simple multiplexing), transformed to binary, and then digitally corrected (simple digital addition) to give the final $n$-bit output digital word. Notice that only $\mathrm{n}-1$ residue voltages are required to calculate the $n$-bit digital word that represents the sampled input voltage $\mathrm{V}_{\mathrm{in}}$. The following observations are made:

## 1) The residue voltage is held across the feedback capacitor:

In flip-around SC amplifiers, the residue voltage is impressed on the feedback capacitor. This fact suggests the use of the feedback capacitor in the SC amplifier as the sampling capacitor for the residue voltage. Therefore, no sampling capacitance is needed at the output of the SC amplifier. Amplification of the output voltage by two can take place by simply switching half of the feedback capacitor to form the flip-around structure [12].

## 2) The final residue amplification is not required:

In the conventional cyclic operation shown in Fig. 3, the last residue voltage generated by $\mathrm{CC}_{\mathrm{a}}$, $\mathrm{CC2}_{\mathrm{b}}$ and the opamp in the nth cycle is not utilized, because $\mathrm{CCl}_{\mathrm{a}}, \mathrm{CCl}_{\mathrm{b}}$, are sampling the input voltage. This suggests the idea of using $\mathrm{CC} 2_{a}, \mathrm{CC}_{\mathrm{b}}$ as well as the opamp for other purposes. If needed, the opamp can be configured for offset cancellation. A new structure, motivated by these two observations, is presented the following section.

## Proposed cyclic structure

The proposed SC structure's states and their sequence, which implement the same recursive function of (1), are shown in Fig. 4. In the 'Initial State' the capacitor pair C2a, C2b and the capacitor pair $\mathrm{C} 1 \mathrm{a}, \mathrm{C} 1 \mathrm{~b}$ both sample the input voltage. At the start of the next state 'State X ', the pair $\mathrm{C} 2 \mathrm{a}, \mathrm{C} 2 \mathrm{~b}$ are switched together to form one feedback capacitance for the first residue amplification, while connecting the $\mathrm{Cla}, \mathrm{Clb}$ pair to $\mathrm{DAC1}$ voltage which is by the comparison result of the input voltage sampled in the 'Initial State'. Since the pair C2a, C2b hold the value of the residue voltage when in 'State X ', amplification during 'State B' can take place by simply connecting C2a to the output of the DAC2 voltage which is determined by the comparison result of the residue voltage of 'State X '.


Figure 4. Switching sequence of the SC networks in the proposed technique

Since the capacitor pair $\mathrm{Cl}_{\mathrm{a}}, \mathrm{Cl}_{\mathrm{b}}$ is not needed for amplification during 'State B ', these capacitors can be used to sample the residue voltage at the amplifier output. This process can then be repeated by alternating between 'State B' and 'State A' until the end of the conversion is reached. The SC networks are then switched back to the 'Initial State' from 'State A' to start a new conversion. The final SC networks for the proposed structure along with the required clocks are shown in Fig. 5. The difference between the proposed structure and the conventional structure is in the elimination of the extra sampling capacitance during the second step (the first 'State A' in the conventional case and 'State $X$ ' in the proposed structure). This difference has a significant impact on overall performance, as will be discussed in the following section.

## A. Benefits of the proposed structure

The performance of a switched capacitor amplifier is dominantly determined by the size of the capacitors and it is this size difference that offers advantages for the proposed circuit shown in Fig. 5. It is well known that the capacitor size in a pipelined ADC becomes increasingly less important as one moves from the MSB stages to the LSB stages in the pipeline. Both reduced matching requirements and reduced effects of $\mathrm{kT} / \mathrm{C}$ noise contribute to this relaxation in requirements. It was observed in [13] that an optimum capacitance stage-scaling factor for a pipeline ADC exists and is approximately equal to reciprocal of the interstage gain. Although aggressive capacitor scaling is practical in a pipelined architecture, capacitor scaling in a cyclic structure becomes temporal rather than spatial and circuit overhead makes it more difficult to take full advantage of capacitor scaling in a cyclic structure. But even in a cyclic structure, significant power and area benefits can be derived with appropriate capacitor sizing and scaling in the first one or two conversion cycles. However, if capacitor scaling is used in a cyclic structure, making the temporal capacitor scaling factor equal to the reciprocal of the interstage gain should give near optimal performance for the cyclic structure as well. Since a 1-bit per clock-phase cyclic structure has a nominal interstage gain of two, good performance should be obtained if the capacitance from one stage to the next is decreased by a factor of 2 . The proposed structure scales the sampling capacitance by a factor of 2 from the 'Initial State' to 'State X '. The same ' $\mathrm{kT} / \mathrm{C}$ ' noise performance at sampling input is maintained when comparing the proposed structure of Fig. 5 with the conventional structure of Fig. 2. This is true as the capacitors in the two circuits will be related by the expressions $\mathrm{C1}_{\mathrm{a}}+\mathrm{Cl}_{\mathrm{b}}=\mathrm{CC}_{\mathrm{a}}$ and $\mathrm{C} 2_{a}+\mathrm{C}_{2}=\mathrm{CC1} 1_{\mathrm{b}}$. Table 1 shows the opamp's capacitive loading in each state for both the conventional circuit and the structure of Fig. 5 assuming that all of the conventional and proposed structures' capacitors are equal to ' 1 C ' and


Figure 5. Proposed SC networks for the cyclic ADC with the required clock signals
' 0.5 C ' respectively, where C is an arbitrary unit. Note that the maximum capacitive load of the proposed structure has been reduced by a factor of 2 from ' 2.5 C ' to ' 1.25 C ' and it is this reduction in total capacitance that not only provides an area reduction for the layout of the capacitors but a significant power reduction in the design of the operational amplifier when the resolution of the cyclic structure is large as well. A fairer comparison may involve scaling the capacitors $\mathrm{CC}_{\mathrm{a}}$ and $\mathrm{CC}_{\mathrm{b}}$ to ' 0.5 C ' each. However, even with this scaling, the opamp will still be required to drive a maximum load of ' 2.25 C ' in 'State B' for the conventional sequence of Fig. 3. On the other hand, the opamp in the proposed structure would still only need to drive a maximum load of ' 1.25 C '. Using a first-order opamp model, it can be shown that this results in a $44 \%$ reduction in this dominant power-consuming component. It is more common to have all of the capacitors of the conventional implementation equal, and hence a $50 \%$ power reduction is more realistic. The energy savings does depend on the architecture of the opamp and circuit parasitics such as the ON-resistance of the switches and parasitic capacitances. If these effects are included, the energy savings will be reduced but significant benefits would still be obtained. Unlike the pipeline structure in [12], the proposed 2-stage cyclic structure samples the input voltage on the same set of capacitors at start of every conversion. Hence, every conversion will suffer consistent gain errors and therefore the proposed structure of Fig. 5 will not introduce a new source of harmonic distortion as in [12].

## B. Design Issues

In the proposed SC switching sequence, 'State X ' introduces two series switches in the signal path in the SC network. This suggests an increase in the time constant of the opamp in this state. However, as indicated by Table 1 , the loading of the opamp is reduced by a factor of five over the conventional implementation and this compensates for this effect well enough to make settling of 'State X ' in the implemented ADC faster than 'State A' and 'State B'. Another design issue was the switching complexity of the proposed structure. However, as shown in Fig. 5, dummy switches were added to simplify clocking while having better layout matching for the capacitors.

Table 1 Capacitive loading comparison

| Structure | Initial State | State X | State B | State A |
| :--- | :--- | :--- | :--- | :--- |
| Conventional | 0.5 C | N/A | 2.5 C | 2.5 C |
| Proposed | 0 | 0.5 C | 1.25 C | 1.25 C |

## Implementation and simulation results

The proposed structure was used in the design of a $10-$ bit, 2.3 MHz cyclic ADC in a $0.5 \mu \mathrm{CMOS}$ process. The opamp architecture is a cascode-cascade structure for high gain and large signal swing. The comparators are dynamic comparators for lower power consumption. The sampling switches are bootstrapped to accommodate the input signal swing. Spectral simulation results are shown in Fig. 6. Although not shown, the ADC demonstrated complete 10-bit performance for low frequencies and for frequencies up to near the Nyquist rate. The thermal noise and mismatches of circuit devices are not incorporated in the simulation results presented.


Figure 6. Reconstructed simulation spectrum of 100 KHz sinusoidal input signal

## Conclusion

A new cyclic ADC structure that is built upon a conventional two stage RSD cyclic ADC with a shared opamp was introduced. The new structure can reduce the total capacitance by up to $50 \%$ and reduce the capacitive loading on the opamp by $50 \%$ as well thus resulting in a reduction of the opamp power consumption and a reduction in the area needed for the capacitor layout. This reduction in
capacitance and power was achieved while maintaining the same SNR and conversion speed performance of the conventional implementation.

## References

[8] P. W. Li, M.J. Chin, P. R. Gray, R. Castello, "Ratio-independent algorithmic analog-to-digital conversion technique", IEEE Journal of Solid-State-Circuits, vol. 19, pp. 828-835, Dec. 1984.
[9] B. Ginetti, P. Jespers, A. Vandemeulebroecke, "A CMOS 13 bits cyclic RSD A/D converter", Proceedings ESSIRC 1991, Milan, pp. 345-348, September 1991.
[10] D. Garrity and P. Rakers, "A $10 \mathrm{bit}, 2 \mathrm{Ms} / \mathrm{s}, 15 \mathrm{~mW}$ BiCMOS cyclic RSD A/D converter", Proceedings of the Bipolar/BiCMOS Circuits and Technology Meeting, pp. 192-195, 1996.
[11] D. Garrity and P. Rakers, "Low power cyclic A/D converter", U.S. Patent 6,535,157, March 18, 2003.
[12] S. Q. Malik and R. L. Geiger, "Simultaneous Capacitor Sharing and Scaling for Reduced Power in Pipeline ADCs", Proceedings of IEEE Midwest Symposium on Circuits and Systems, August 2005.
[13] D. W. Cline, and P. R. Gray, "A power optimized 13-b 5 Msamples/s pipelined analog-to-digital converter in $1.2 \mu \mathrm{~m}$ CMOS", IEEE Journal of Solid-State Circuits, vol. 31, pp. 294-303, March 1996.

# CHAPTER 4. A LOW TEMPERATURE SENSITIVITY SWITCHED-CAPACITOR CURRENT REFERENCE 

A paper published in the Proceedings of the<br>2001 European Conference on Circuit Theory and Design ${ }^{1}$

S. Q. Malik, M. E. Schlarmann, and R. L. Geiger


#### Abstract

A current reference with low temperature sensitivity based on a switched-capacitor technique has been developed. The implementation is targeted for a $0.18 \mu$ CMOS process. HSPICE simulations using level 49 models valid over a wide temperature range were used to verify the design. The simulation results predict variations of less than $\mathbf{0 . 0 2 9} \%$ over a temperature range of $-\mathbf{4 0}{ }^{\circ} \mathrm{C}$ to $125^{\circ} \mathrm{C}$.


## Introduction

Current references are needed in many analog signal processing applications including operational amplifier (opamp) and data converter bias circuits. These applications often require a reference current with low temperature dependence.

Unlike voltage references that can be derived from intrinsic physical values of the process, no intrinsic current reference is available in CMOS [1]. As a result, reference currents are often obtained by applying a temperature stable voltage (obtained from a voltage reference) across a resistor. The resistor is either integrated on-chip or may be supplied off-chip for improved control over temperature characteristics. However, both cases have drawbacks. On-chip resistors typically exhibit large temperature dependence while off-chip resistors are often not a feasible option for many applications due to cost and area considerations. This work circumvents the need for an accurate onchip resistor by using a switched capacitor technique to generate a temperature independent current.

The previous work in the area is briefly surveyed in section 2 . The newly proposed structure is introduced in section 3. Design considerations and modifications to handle certain requirements are detailed in section 4 while simulation results are presented in section 5 followed by conclusions. ${ }^{1}$

[^0]
## Background

Approaches that use a resistor to generate a temperature independent current reference have been reported in [1]-[3]. Due to the large temperature coefficients of polysilicon and well diffusions, monolithic resistors exhibit large temperature dependence. To overcome this problem, resistorless architectures have also been developed [4][5]. Integrated capacitors can be fabricated with greater precision and exhibit significantly lower temperature dependence than integrated resistors. Therefore, switched capacitor methods of generating temperature stable currents have emerged [6]-[8]. This paper presents a current reference using switched-capacitor based circuit to deliver and maintain a stable current.

## Current reference architecture

Precise crystal-based clocks and temperature independent voltage references are commonly available on-chip. Given that fact, a temperature stable current can be developed using a switched capacitor technique. The concept involves periodically dumping a fixed amount of charge onto a circuit node whose time-average value is held fixed by a feedback network.

The circuit operates as follows. $\phi_{1}$ and $\phi_{2}$ are non-overlapping clocks of frequency $f_{\text {clk }}$. The amplifier is assumed to have a single pole response. Its speed (unity gain frequency) is intentionally set very


Figure 1: Proposed current reference
low so that it is fast enough to respond to temperature variations yet slow enough to be unable to effectively respond to signals operating at the clock frequency. During $\phi_{1}, \mathrm{C}_{1}$ charges to $\mathrm{V}_{\text {ref }}$. During $\phi_{2}$ the charge on $\mathrm{C}_{1}$ is dumped onto node 1 . The instantaneous change in voltage on node $\mathbf{1}$ due to this charge is given by

$$
\begin{equation*}
\Delta V=-\left(\frac{C_{1}}{C_{1}+C_{2}}\right) \cdot\left(V_{r e f}+V_{1}\right) \tag{1}
\end{equation*}
$$

where $\mathrm{V}_{1}$ is the voltage on node 1 immediately preceding the charge transfer, as shown in Fig. 2. The amplifier responds to the low-frequency component of the signal on node 1 . Over time, it adjusts the bias on $\mathrm{M}_{1}$ so that the time average value of the signal present on node 1 is zero. In steady state, the signal on node 1 looks like the one shown in Fig. 2. It is a sawtooth type waveform centered about zero. The sawtooth shape arises due to the steady charging of $M_{1}$ interrupted by the periodic charge transfers from the switched capacitor network. The peak-to-peak magnitude of the signal is given by (1). Since it is centered about zero, the voltage on node 1 just prior to the charge transfer, $\mathrm{V}_{1}$, is approximately given by

$$
\begin{equation*}
V_{1}=-\frac{\Delta V}{2} \tag{2}
\end{equation*}
$$



Figure 2: Voltage at node 1 of Fig. 1

Substituting (2) into (1) and solving for $\Delta \mathrm{V}$ yields the peak-to-peak magnitude of the ripple on node 1 in terms of fixed parameters. This $\Delta \mathrm{V}$ is given by

$$
\begin{equation*}
\Delta V=-\left(\frac{2 C_{1}}{C_{1}+2 C_{2}}\right) \cdot V_{r e f} \tag{3}
\end{equation*}
$$

Thus, the ripple on node 1 can be controlled via the ratio $\mathrm{C}_{2} / \mathrm{C}_{1}$. A large $\mathrm{C}_{2}$ results in reduced ripple at the expense of increased die area. The current delivered by the switched capacitor network is approximately given by

$$
\begin{equation*}
I_{r e f} \cong C_{1} \cdot V_{r e f} \cdot f_{c l k} \tag{4}
\end{equation*}
$$

Due to process variability, the actual post-fabrication value of the current can exhibit significant deviation from the designed value. However, due to the low temperature coefficient of monolithic capacitors, for a given die the current should remain relatively constant over temperature variations.

## Design considerations

Design choices affect the transient startup time, the amount of output ripple, and the stability in presence of a temperature dependent load. To help the designer make intelligent tradeoffs, each of these issues is discussed in this section.

## Increasing the output resistance

Due to the finite output impedance of $\mathrm{M}_{2}$, some temperature dependence is introduced if the drain voltage of $\mathrm{M}_{2}$ is allowed to vary. This is an especially important issue if the load is temperature dependent. To address this issue, output impedance enhancement may be required. One possible method, the regulated cascode, is shown in Fig. 3. In less sensitive situations, standard cascoding may suffice. Note that in Fig. 3, drain voltage of $\mathrm{M}_{2}$ is fixed at 0 V thereby facilitating accurate current mirroring of the reference current. Cascoding not only improves the output resistance of the current reference but also reduces the sensitivity to supply voltage variations.

## Hold capacitor and ripple

Since the amplifier is intentionally made slow, it attenuates the high-frequency components of the signal present on node 1 . However, its response is not zero at those frequencies. Consequently, some ripple will be present in the output current. Fortunately, the magnitude of the ripple can be managed


Figure 3: Proposed circuit with improved output resistance


Figure 4: Proposed circuit with filter


Figure 5: A simple 2nd order filter using MOSCAPs
by controlling the ratio $\mathrm{C}_{2} / \mathrm{C}_{1}$ and the gain-bandwidth product of the amplifier. Reducing the gain bandwidth product of the amplifier will result in less output current ripple but it will also affect how fast the system will respond to temperature changes. Since temperature changes are generally low frequency in nature, reducing the speed of the amplifier is acceptable but it will extend the length of the transient startup period.

For applications with very low ripple requirements, a filter can be inserted as shown in Fig. 4. A simple filter such as one shown in Fig. 5 can be used. Since precise filter characteristics are not required in this application, capacitors can be implemented as MOSCAPs [9] and resistors can be implemented using triode region transistors.

## Improving the settling time

As previously mentioned, the opamp was intentionally made slow in order to reduce the ripple present in the output current. The inevitable consequence of this choice is a longer time for the output of the opamp to settle to its final value. The long settling time has a major impact on the amount of time it takes for the circuit to start up. Once locked to its final value, the output should track slow changes in temperature.

In applications that require faster startup, the proposed circuit can be modified to achieve that without increasing the output ripple by including the filter (as shown in Fig. 4) and increasing the gain-bandwidth product of the amplifier.

## Simulation results

The circuit of Fig. 1 was simulated using HSPICE with level 49 models for a $0.18 \mu$ CMOS process. The models were valid from $-40^{\circ} \mathrm{C}$ to $125{ }^{\circ} \mathrm{C}$. By using a clock frequency of 20 MHz , a $\mathrm{V}_{\text {ref }}$ of $1.25+\mathrm{V}_{\mathrm{ss}}$, and a $\mathrm{C}_{1}$ of 0.25 pF , a reference current $\mathrm{I}_{\text {ref }}$ of $6.25 \mu \mathrm{~A}$ was expected. Since a single-pole


Figure 6: Average output current vs. Temperature
behavioral model was used to model the opamp, the possible temperature dependence of the amplifier is not represented in the results. Furthermore, since models for the temperature variation of poly-poly or metal-metal capacitors were not available, $\mathrm{C}_{1}$ was modeled as temperature independent. However, the temperature dependence of these capacitors is expected to be small in practice.

The circuit was simulated at several points over a temperature range from $-40{ }^{\circ} \mathrm{C}$ to $125{ }^{\circ} \mathrm{C}$. As shown in Fig. 6, the current is very stable over the entire temperature range. The maximum deviation from the midpoint current value is $0.029 \%$.

The actual value of the current obtained was approximately $6.88 \mu \mathrm{~A}$ instead of $6.25 \mu \mathrm{~A}$. The reason for this discrepancy is the non-ideal nature of the virtual ground established at node 1 of Fig. 1. As shown in Fig. 2, the voltage at node 1 is non-zero despite having an approximate average value of 0 . As mentioned in section 4.2 , increasing the size of capacitor $C_{2}$, i.e., the ratio $C_{2} / C_{1}$, can reduce the voltage change on node 1 . As the ripple on node 1 becomes smaller, a more accurate charge transfer from $\mathrm{C}_{1}$ to $\mathrm{C}_{2}$ takes place and the reference current approaches its intended value.

## Conclusions

A new temperature stable current reference was developed. The proposed circuit uses switched capacitor technique to establish the reference current. Simulation results show that the output current varies less than $0.029 \%$ over a temperature range of $-40^{\circ} \mathrm{C}$ to $125^{\circ} \mathrm{C}$.

## Acknowledgements

Support for this project was provided, in part, by the R. J. Carver Trust and National Semiconductor Inc.

## References

[1] E. Vittoz, "The design of high performance analog circuits on digital CMOS chips," IEEE J. Solid-State Circuits, vol. SC-20, no. 3, pp 657-665, June 1985.
[2] C. H.- Lee and H. J.- Park, "All-CMOS temperature independent current reference," Electronics Letters, vol. 32, no. 14, July 1996.
[3] E. Vittoz and J. Fellrath, "CMOS analog circuits based on weak inversion operation," IEEE J. Solid-State Circuits, vol. SC-12, pp 224-231, June 1977.
[4] W. M. Sansen, F. O. Eynde, and M. Steyaert, "A CMOS temperature-compensated current reference," IEEE J. Solid-State Circuits, vol. 23, pp 821-824, June 1988.
[5] H. J. Oguey and D. Aebischer, "CMOS current reference without resistance," IEEE J. Solid-State Circuits, vol. 32, no. 7, pp 1132-1135, July 1997.
[6] A. Olesin, K. L. Luke, and R. D. Lee, "Switched capacitor precision current source," U.S. Patent, 4,374,357 Feb. 15, 1983. Available WWW: http://www.uspto.gov.
[7] G. Torelli and A. D. LaPlaza, "Tracking switched-capacitor CMOS current reference," IEE Proc.Circuits Devices Syst., vol. 145, no. 1, February 1998.
[8] R. H. Leonowich, "Switched capacitor current reference," U.S. Patent, 5,408,174 April 18, 1995. Available WWW: http://www.uspto.gov.
[9] H. Yoshizawa, Y. Huang, P. F. Ferguson, Jr., and G. C. Temes, "MOSFET-only switchedcapacitor circuits in digital CMOS technology," IEEE J. Solid-State Circuits, vol. 34, no. 6, pp 734-747, June 1999.

# CHAPTER 5. AREA EFFICIENT LAYOUT STRATEGIES FOR EXTREME-RATIO MOS TRANSISTORS 

To be submitted<br>Saqib Q. Malik and Randall L. Geiger


#### Abstract

Several different layout schemes that are useful for implementing extreme-ratio low resistance switches with MOS transistors are discussed and characterized. A comparison of the area required for implementing a switch with a standard alternating bar approach is made with layouts using waffle structures, zipper structures, hexagonal structures, and new modified waffle structures. Simple analytical design equations for these non-conventional geometries are introduced. Comparisons show that in typical processes, area reductions of over $40 \%$ are readily achievable with the modified waffle structures.


## Introduction

The effective resistance of MOS transistor operated as a switch is characterized by several parameters. The four that generally receive the most attention are the transistor width-to-length ratio, W/L, the excess bias, the series diffusion resistance, and the contact resistance. For MOS transistors used as switches that must achieve extremely low on-resistances, large effective W/L ratios are used along with multiple contacts to the drain and source diffusions. We refer to such transistors as extreme-ratio devices.

For most layouts of such devices, the total resistance of the switch can be expressed as the sum of three resistances. One termed $\mathrm{R}_{\mathrm{FET}}$, represents the "on" resistance associated with the channel of the transistor itself and is determined by the effective W/L ratio of the MOSFET and the excess bias voltage. A second, termed $\mathrm{R}_{\text {via }}$, is due to the contact resistance to the drain and source diffusions of the switch. The third, termed $\mathrm{R}_{\text {diff }}$, is due to the series resistance in the diffusions between the edge of the channel and the contacts. The resistance associated with the metal interconnects is generally negligible compared to these three resistances. Thus, a single-transistor MOS switch can be modeled by a resistor expressed as

$$
\begin{equation*}
R_{s w}=R_{F E T}+R_{v i a}+R_{d i f f} \tag{1}
\end{equation*}
$$



Fig. 1 (a) Typical Layout, (b) An irregular transistor

For the simple MOS switch driven on with a control voltage of $V_{D D}$ and with the layout shown in Fig. 1a, it follows from the simple square-law device model that the three parts are approximately given by

$$
\begin{gather*}
R_{F E T}=\frac{1}{\mu C_{O X} \cdot \frac{W}{L} \cdot\left(V_{D D}-V_{T}\right)}  \tag{2}\\
R_{v i a}=2 \cdot R_{c o n t}  \tag{3}\\
R_{d i f f}=R_{S q} \cdot \frac{a+b}{W} \tag{4}
\end{gather*}
$$

where $\mu$ is the mobility of carriers in the channel, $C_{O X}$ is the gate oxide capacitance density, $V_{T}$ is the threshold voltage of the MOSFET, $\mathrm{R}_{\text {cont }}$ is the contact resistance, and $\mathrm{R}_{\mathrm{sq}}$ is the diffusion sheet resistance. For a near minimum-sized n-channel transistor in AMI $0.5 \mu$ process with $\mathrm{W}=\mathrm{L}, \mathrm{a}=\mathbf{b}=$ $2 \cdot \mathrm{~W}, \mathrm{R}_{\mathrm{sq}}=5 \Omega / \square, \mu \mathrm{C}_{\mathrm{Ox}}=118 \mu \mathrm{~A} / \mathrm{V}^{2}, \mathrm{~V}_{\mathrm{T}}=0.8 \mathrm{~V}, \mathrm{~V}_{\mathrm{DD}}=3.3 \mathrm{~V}$ and $\mathrm{R}_{\mathrm{cont}}=55 \Omega$, the total switch resistance as comprised by the three parts of (1) becomes:

$$
\begin{equation*}
R_{s w}=3390+110+160=3660 \Omega \tag{5}
\end{equation*}
$$

The contributions due to $\mathrm{R}_{\text {via }}$ and $\mathrm{R}_{\text {diff }}$ thus represent about $3 \%$ and $4.4 \%$ of the total resistance respectively. It can be concluded that for this simple structure, the contact resistances and the diffusion resistances are negligible. On structures with very large $\mathrm{W} / \mathrm{L}$ ratios, the term $\mathrm{R}_{\text {FET }}$ can be driven to an arbitrarily low value. Correspondingly, with most common layout schemes, multiple vias will be made to contact the diffusions thus driving $\mathrm{R}_{\text {via }}$ and $\mathrm{R}_{\text {diff }}$ down as well. With a little care in layout, these two resistances will scale approximately linearly with $\mathrm{R}_{\mathrm{FET}}$ thus keeping their contribution to $\mathrm{R}_{\mathrm{sw}}$ negligible. For this reason, throughout this paper, the contributions to the total resistance due to the last two terms in (1) will be neglected.

Some applications require switch resistance in the few ohms range or even smaller. From (2), it is apparent that extreme (very large) effective $\mathrm{W} / \mathrm{L}$ ratios are required to achieve this. For example, a switch with an on-resistance of $1 \Omega$ would require an effective W/L ratio of about 3400 in the typical process referenced above. The silicon area implications associated with such an extreme-ratio transistor are significant. Correspondingly, a MOS transistor with an actual rectangular layout and an aspect ratio of $3400: 1$ would not be practical. The large effective W/L ratio is generally achieved by using a serpentined layout structure in which the MOS transistor is folded to make the aspect ratio of the footprint of the MOS transistor reasonable or even nearly square. A standard variant of the serpentine that can further reduce the area of the overall transistor footprint is based upon using alternating bars of source and drain diffusion that are shared and interconnected. Such a layout scheme is shown in Figure 2.


Fig. 2: (a) Alternating bar, (b) Reference cell

Significant improvement in area efficiency over what is achievable by the serpentined structure or the alternating bar structure is possible by judicious selection of alternative but less popular layout schemes. These alternative layout schemes can offer economic benefits where substantial portion of a design is devoted to switches that must have low on resistance. In addition, these alternate layout schemes typically have less gate capacitance for a given effective W/L ratio. This reduces the capacitive loading and correspondingly the speed of circuits driving these switches while simultaneously reducing the dynamic power dissipation in the switches. Several alternative layout schemes will be discussed after developing a method for comparing alternative layout structures.

## Layout comparison method

Most layouts of extreme-ratio transistors are based upon attempts to replicate not just the large W/L ratio of the transistor but also the rectangular aspect ratios of the transistor. The layout methods discussed here will not be based upon any attempt to preserve a rectangular gate region for the transistor. It has been shown [1] that corresponding to any arbitrary shaped device that has two disconnected diffusion regions separated by a channel region, there is a rectangular MOS transistor that has the same dc I-V characteristics. This equivalence is depicted in Fig. 1 where the dark region in Fig. lb denotes the channel region of an arbitrarily-shaped transistor, $d_{l}$ and $s_{l}$ denote the disconnected diffusion regions that serve as the drain and source of the arbitrarily-shaped transistor, and $d_{l}$ and $s_{l}$ denote the corresponding drain and source regions of an equivalent rectangular transistor. For notational convenience, we will refer to the $\mathrm{W} / \mathrm{L}$ ratio of a corresponding rectangular transistor as the "effective W/L ratio", (W/L) eff, of the arbitrarily shaped transistor. In the proof of the existence of the rectangular transistor with an equivalent $\mathrm{W} / \mathrm{L}$ ratio, another useful result that will be used later in this paper was obtained. Specifically, since the regions denoted by $\mathrm{d}_{2}$ and $\mathrm{s}_{2}$ in Fig lb are also disconnected, they can likewise be used to form the drain and source of a second arbitrarilyshaped transistor which we will term the reciprocal transistor to the original device. The same theorem thus guarantees a second equivalent rectangular transistor and it was shown that this transistor has an effective W/L ratio that is the reciprocal to that of the rectangular transistor in Fig. lb. Thus, the reciprocal transistor for the original arbitrarily-shaped transistor has as an equivalent rectangular transistor that is the reciprocal transistor for the original rectangular transistor. This is also depicted in Fig. 1. Thus, the effective W/L ratio of a reciprocal transistor is the reciprocal of the W/L ratio of the original transistor.

In what follows, it will be seen that there are substantial area benefits associated with non-rectangular transistors when large effective $\mathrm{W} / \mathrm{L}$ ratios are required. The effective utilization of layouts that are not based upon rectangular transistors requires a systematic procedure for determining the effective W/L ratio of nonrectangular structures along with the corresponding area. If the periphery effects are neglected, we will see in what follows that all extreme-ratio structures that will be considered can be represented by the parallel interconnection of an arbitrary number of smaller structures. These smaller structures will be referred to as reference cells. We will characterize the effective $\mathrm{W} / \mathrm{L}$ ratio and the area of the reference cells and then extend these results to obtain the effective $\mathrm{W} / \mathrm{L}$ ratio for the overall structure by connecting n of these references cells in parallel.

If $R_{\text {des }}$ is the maximum acceptable resistance of a switch then, if via and diffusion resistances are neglected and it is assumed we are interested in extreme ratio switches so that the periphery effects of the cell are negligible, the number of reference cells, $n$, and the area needed to achieve this resistance are given by the expressions

$$
\begin{gather*}
n=\operatorname{int}_{L}\left(\frac{R_{r e f}}{R_{d e s}}\right) \\
A=\operatorname{int}_{L}\left(\frac{R_{r e f} \cdot A_{r e f}}{R_{\text {des }}}\right) \tag{6}
\end{gather*}
$$

where the function $\operatorname{int}_{\mathrm{L}}(\mathrm{x})$ denotes the smallest integer greater than $\mathrm{x}, \mathrm{R}_{\mathrm{ref}}$ is the resistance of the reference cell and $\mathrm{A}_{\text {ref }}$ is the area of the reference cell which includes any drain/source diffusions and interconnect spacing needed for the reference cell. Since we are interested in large ratios of $R_{\text {ref }} / R_{\text {des }}$,

$$
\begin{gather*}
\mathrm{n} \approx\left(\frac{\mathrm{R}_{\mathrm{ref}}}{\mathrm{R}_{\mathrm{des}}}\right) \\
A \approx \frac{R_{r e f} \cdot A_{r e f}}{R_{d e s}} \tag{6a}
\end{gather*}
$$

An accurate determination of the resistance of a switch is strongly dependent upon the model used for the device. Although it is often argued that the square-law model is not adequate for accurately predicting the resistance of a switch, the relative value of the resistance of a rectangular device and the resistance of a nonrectangular device is not strongly dependent upon the difference between the square law model and the much more complicated and widely-used BSIM model. Thus, for notational convenience, throughout this paper it will be assumed that the square-law device model can be used
when a transistor is operated deep in the triode region. It thus follows from the square-law device model that the resistance of the reference cell is given by the expression

$$
\begin{equation*}
R_{r e f}=\frac{1}{\mu C_{o x} \cdot\left(\frac{W}{L}\right)_{e f f} \cdot\left(V_{D D}-V_{T}\right)} \tag{7}
\end{equation*}
$$

where (W/L) $)_{\text {eff }}$ is the effective W/L ratio of the reference cell. From (6a) and (7), it follows that the normalized reference area, $\mathrm{A}_{\text {ret,n }}$, defined by

$$
\begin{equation*}
A_{r e f . n}=\frac{A_{r e f}}{\left(\frac{W}{L}\right)_{e f f}} \tag{8}
\end{equation*}
$$

is a figure of merit for comparing the area efficiency of different layout structures since the total area is proportional to $\mathrm{A}_{\text {ref.n. }}$ Reference cells with smaller values of $\mathrm{A}_{\text {ref.. }}$, will require less total area than reference cells with larger values for this metric.

Knowledge of the capacitances associated with the transistor for a specific layout structure is also important for complete characterization of the structure. This information is particularly useful since existing extraction tools will not be able to accurately extract the parasitic capacitances in highly irregular transistor structures. A good approximation to these capacitances can be obtained by knowing the actual gate area, the actual drain and source area, and the actual drain and source perimeter. Parameterized expressions that are applicable in an arbitrary technology for (W/L) eff, $\mathrm{A}_{\mathrm{ref}}$, $\mathrm{A}_{\text {ref. }}$, as well as the gate area and the diffusion area and diffusion perimeter for the drain and source diffusions of the reference cell for several different area-efficient layouts will be presented in the next section. The corresponding total gate capacitance as well as the drain and source diffusion capacitances for large-ratio switches can be obtained by multiplying the appropriate derived values for the reference cell by $n$.

## Layout structures

Several different layout structures will be discussed in this section. The switch area efficiency for a given layout structure is dependent upon the minimum allowable feature sizes as characterized by the design rules of the process. The design rules that impact area efficiency of a switch layout are denoted
by $d_{l} \ldots d_{7}$ in Table 1 . The typical value for these parameters for $\lambda$-based design rules are also given in the table.

In what follows, dark grey is used to represent polysilicon and light gray to represent diffusion regions. Solid squares are used to denote drain contacts and dotted squares are used to denote source contacts. All polysilicon is assumed connected to form the gate of a transistor. Likewise, all diffusions with source contacts are assumed connected together to form the source of the switch and all drains are connected together to form the drain. For notational convenience, all structures are shown without metal interconnects. In this section, emphasis will be placed on characterizing only the reference cells for a given layout, that is, all periphery effects will be ignored. This assumption is reasonably good for extreme-ratio switches but the periphery effects will play a significant role if extreme ratios are not needed. The periphery effects will be considered in the next section.

Table 1: Typical MOSIS design rules for MOS switch layouts

| Rule (minimum) | Name | $\lambda$-based rule |
| :--- | :--- | :--- |
| Poly Width | $d_{1}$ | $2 \lambda$ |
| Diffusion Width | $d_{2}$ | $3 \lambda$ |
| Contact Opening | $d_{3} \times d_{3}$ | $2 \lambda \times 2 \lambda$ |
| Contact-Poly Spacing | $d_{4}$ | $2 \lambda$ |
| Diffusion Overlap of Contact | $d_{5}$ | $1.5 \lambda$ |
| Contact-Contact Spacing | $d_{6}$ | $2 \lambda$ |
| Poly-Poly Spacing | $d_{7}$ | $2 \lambda$ |

## Alternating Bar structure

The widely used alternating bar structure is shown in Fig. 2a along with a reference cell for this layout. The structure is characterized by the parallel interconnection of multiple instantiations of the reference cell which is expanded in Fig. 2b. The area, (W/L) eff and $\mathrm{A}_{\text {ref.n }}$ for this reference cell are given respectively in terms of the design rules by:

$$
\begin{gather*}
A_{r e f 1}=2\left(d_{3}+d_{6}\right) \cdot\left(d_{1}+d_{3}+2 d_{4}\right)  \tag{9}\\
\left(\frac{W}{L}\right)_{e f f 1}=2 \cdot \frac{d_{3}+d_{6}}{d_{1}}  \tag{10}\\
A_{r e f, n 1}=d_{1} \cdot\left(d_{1}+d_{3}+2 d_{4}\right) \tag{11}
\end{gather*}
$$

For the alternating bars reference cell, the gate area and the drain and source diffusion areas are given by

$$
\begin{gather*}
A_{\text {gatel }}=2 \cdot d_{1}\left(d_{3}+d_{6}\right)  \tag{12}\\
A_{\text {diff } 1}=2\left(d_{3}+d_{6}\right)\left(d_{3}+2 d_{4}\right) \tag{13}
\end{gather*}
$$

To find the perimeter of the diffusion region for a given reference cell, we will assume that the reference cell is surrounded by similar reference cell. The diffusion perimeter is then given by

$$
\begin{equation*}
P_{\text {diff } 1}=4\left(d_{3}+d_{6}\right) \tag{14}
\end{equation*}
$$

## Waffle structure

The waffle structure is shown in Fig. 3a. with a reference cell shown in the center of and expanded in Fig. 3b. This structure is well known [2]-[8] and is similar to the structure used in vertical power MOSFETS [3]. For this reference cell, (W/L) eff is not readily attainable directly but from results in Section II, the equivalent W/L ratio of the reciprocal transistor can be obtained and by taking the reciprocal of this, the W/L ratio of the desired transistor is attained. In obtaining the W/L ratio of the reciprocal transistor, a $90^{\circ}$ angle in the channel is encountered. As is often done when calculating the number of squares in a resistive region, the approximation of using $0.55 \mathrm{~W} / \mathrm{L}$ units for the $90^{\circ}$ bend in


Fig. 3: (a) Waffle structure

(b) Reference cell
the channel was used. The effective area and an estimate of $(\mathrm{W} / \mathrm{L})_{\text {eff }}$ for this reference cell are given respectively by:

$$
\begin{gather*}
A_{\text {ref } 2}=2 \cdot\left(d_{1}+d_{3}+2 d_{4}\right)^{2}  \tag{15}\\
\left(\frac{W}{L}\right)_{e f f 2}=2 \cdot \frac{2\left(d_{3}+2 d_{4}\right)+0.55 d_{1}}{d_{1}} \tag{16}
\end{gather*}
$$

The normalized area of the reference cell is given by

$$
\begin{equation*}
A_{r e f, n 2}=\frac{d_{1} \cdot\left(d_{1}+d_{3}+2 d_{4}\right)^{2}}{2\left(d_{3}+2 d_{4}\right)+0.55 d_{1}} \tag{17}
\end{equation*}
$$

Similarly, the gate area, and the diffusion area and perimeter of the drain and source regions for the waffle reference cell are given by

$$
\begin{gather*}
A_{\text {gate } 2}=2 \cdot d_{1}\left(d_{1}+2 d_{3}+4 d_{4}\right)  \tag{18}\\
A_{d i f f 2}=2\left(d_{3}+2 d_{4}\right)^{2}  \tag{19}\\
P_{\text {diff } 2}=8\left(d_{3}+2 d_{4}\right) \tag{20}
\end{gather*}
$$

## Zipper structure

The Zipper structure is shown in Fig. 4a. The reference cell depicted in the center of Fig. 4a.is expanded in Fig. 4b. For this reference cell, (W/L) eff is not readily attainable directly but using the reciprocal transistor approach as outlined earlier, the area, an estimate of (W/L) eff, and $\mathrm{A}_{\text {ref.a }}$ for the reference cell can be readily obtained. They are given respectively by:

$$
\begin{gather*}
A_{r e f 3}=2 \cdot\left(d_{1}+d_{2}\right)\left((x+3) \cdot d_{1}+d_{3}+2 d_{4}\right)  \tag{21}\\
\left(\frac{W}{L}\right)_{e f 3}=2 \cdot \frac{\left(2.1 \cdot d_{1}+d_{2}+x \cdot d_{1}\right)}{d_{1}}  \tag{22}\\
A_{r e f \cdot n 3}=\frac{d_{1} \cdot\left(d_{1}+d_{2}\right)\left((x+3) \cdot d_{1}+d_{3}+2 d_{4}\right)}{2.1 \cdot d_{1}+d_{2}+x \cdot d_{1}} \tag{23}
\end{gather*}
$$

The parameter $x$ shown in Fig. 4 b is a degree of freedom that can be used to characterize the depth of the fingers. For $x=0$, (23) reduces to


$$
\left.A_{\text {ref. } n 3}\right|_{x=0}=\frac{d_{1} \cdot\left(d_{1}+d_{2}\right)\left(3 d_{1}+d_{3}+2 d_{4}\right)}{2.1 \cdot d_{1}+d_{2}}
$$

The deep zipper structure is obtained by increasing the depth of the fingers to $x=2$. With this change, we obtain

$$
\begin{equation*}
\left.A_{r e f \cdot n 3}\right|_{x=2}=\frac{d_{1} \cdot\left(d_{1}+d_{2}\right)\left(5 d_{1}+d_{3}+2 d_{4}\right)}{4.1 \cdot d_{1}+d_{2}} \tag{24}
\end{equation*}
$$

The infinitely deep zipper structure, obtained by making the fingers arbitrarily long, (i.e., letting $x$ approach infinity) is characterized by

$$
\begin{equation*}
\lim _{x \rightarrow \infty} A_{r e f, n 3}=d_{1} \cdot\left(d_{1}+d_{2}\right) \tag{25}
\end{equation*}
$$

From a practical viewpoint, the depth of the fingers is limited since deep fingers will add considerable series resistance in the drain and source regions. Depending on a given process, when $x$ gets much beyond 2 or 3 , the series diffusion resistance can start to become a significant portion of the overall resistance thus negating further reductions in switch impedance by making the fingers deeper.

The gate area and the diffusion area and perimeter required to find capacitance for the reference cell are respectively given by

$$
\begin{gather*}
A_{\text {gate } 3}=2 \cdot d_{1}\left((x+3) d_{1}+d_{2}\right)  \tag{26}\\
A_{d i f 3}=2 \cdot\left[\left(d_{1}+d_{2}\right)\left(d_{3}+2 d_{4}\right)+2 \cdot(x+2) \cdot d_{1} \cdot d_{2}\right]  \tag{27}\\
P_{\text {diff } 1}=4\left[(x+3) d_{1}+d_{2}\right] \tag{28}
\end{gather*}
$$

## Star Zag

The Star Zag structure is shown in Fig 5a and the expanded reference cell for this structure is shown in Fig.5b., It can be shown that the area, effective W/L, and the normalized area for the star zag structure are given respectively by

$$
\begin{gather*}
A_{r e f 4}=2 \cdot\left(4 d_{1}+2 d_{2}\right) \cdot\left(3 d_{1}+3 d_{2}\right)  \tag{29}\\
\left(\frac{W}{L}\right)_{e f f}=\frac{16.6 d_{1}+10 d_{2}}{d_{1}}  \tag{30}\\
A_{r e f, n 4}=\frac{d_{1} \cdot\left(4 d_{1}+2 d_{2}\right) \cdot\left(3 d_{1}+3 d_{2}\right)}{8.3 d_{1}+5 d_{2}} \tag{31}
\end{gather*}
$$

The gate area and the diffusion area and perimeter for the Star-Zag structure are given by

$$
\begin{gather*}
A_{\text {gale } 4}=2 \cdot d_{1}\left(11 d_{1}+5 d_{2}\right)  \tag{33}\\
A_{\text {diff }}=2\left(\left(d_{1}+2 d_{2}\right)^{2}+d_{2}\left(9 d_{1}+2 d_{2}\right)\right)  \tag{34}\\
P_{\text {diff } 4}=2\left(22 d_{1}+10 d_{2}\right) \tag{35}
\end{gather*}
$$



Fig. 5: (a) Star Zag structure (b) Reference cell

## Fingered-Waffle

The diffusion regions of the waffle structure, shown earlier, can be extended to for "fingers". The resulting "Fingered-Waffle" structure is shown in Fig. 6a and the reference cell for this structure is shown in Fig. 6b. The distance x shown in the reference cell is a variable and can take on any nonnegative value. An exact analysis of this structure is not practical but if we approximate each square of channel in a $90^{\circ}$ corner in the reciprocal device by $0.55 \mathrm{~W} / \mathrm{L}$ units, a good approximation that is modestly smaller than the actual (W/L) eff can be obtained. Approximate the following expressions for the area, effective W/L, and the normalized area (for $\mathrm{x} \geq 1$ ) can be obtained:

$$
\begin{gather*}
A_{\text {ref } 5}=2 \cdot\left[\left((1+x) d_{1}+d_{3}+2 d_{4}\right) \cdot\left(2 d_{1}+2 d_{2}\right)\right]  \tag{37}\\
\left(\frac{W}{L}\right)_{\text {eff } 5}=\frac{2\left[(2 x-1) d_{1}+2 d_{2}+d_{3}+2 d_{4}+3 \cdot 0.55 d_{1}\right]}{d_{1}}  \tag{38}\\
A_{\text {refn5 }}=\frac{d_{1} \cdot\left((1+x) d_{1}+d_{3}+2 d_{4}\right) \cdot\left(2 d_{1}+2 d_{2}\right)}{(2 x-1) d_{1}+2 d_{2}+d_{3}+2 d_{4}+3 \cdot 0.55 \cdot d_{1}} \tag{39}
\end{gather*}
$$

For the fingered-waffle reference cell, the gate area and the diffusion area and perimeter are given by

$$
\begin{gather*}
A_{\text {gate } 5}=2 \cdot d_{1}\left(4(1+x) d_{1}+3\left(d_{3}+2 d_{4}\right)+4 d_{2}\right)  \tag{40}\\
A_{\text {diff } 5}=2 d_{2}\left(2 x \cdot d_{1}+2 d_{3}+4 d_{4}\right)+d_{1}\left(d_{3}+2 d_{4}\right)  \tag{41}\\
P_{\text {diff } 5}=2 \cdot\left[2(2 x+1) \cdot d_{1}+4 d_{2}+2 d_{3}+4 d_{4}\right] \tag{42}
\end{gather*}
$$



Fig. 6: (a) Fingered-Waffle
(b) Reference cell

## Hexagonal

The hexagonal shaped transistor has been studied recently for matching properties [7] and some have proposed using this to improve area efficiency in extreme-ratio applications. The hexagonal transistor is shown in Fig. 7a and a reference cell for this structure is shown in Fig. 7b. The effective W/L is not easily obtained for this structure. In [7], the effective W/L was derived to be

$$
\begin{equation*}
\left(\frac{W}{L}\right)_{\text {eff } 6 a}=\frac{6}{\ln \left(\frac{W_{2}}{W_{1}}\right) \cdot \cos 30^{\circ}} \tag{43}
\end{equation*}
$$

where $W_{1}$ and $W_{2}$ are the physical parameters shown in Fig. 7b. The effective $\mathrm{W} / \mathrm{L}$ can also be derived using the method described in [1] and can approximated by

$$
\begin{equation*}
\left(\frac{W}{L}\right)_{e f f 6}=6 \cdot\left(\frac{W_{1}}{d_{1}}+0.4\right) \tag{44}
\end{equation*}
$$

The parameters $W_{1}$ and $W_{2}$ can also be expressed in terms of the design rules variables of Table 1. With minimum spacing and square contacts, W1 and W2 are given by

$$
\begin{gather*}
W_{1}=d_{2}+\frac{d_{3}}{2}  \tag{45}\\
W_{2}=\frac{2 d_{1}}{\sqrt{3}}+d_{2}+\frac{d_{3}}{2} \tag{46}
\end{gather*}
$$



Fig. 7: (a) Hexagonal (b) Reference cell

The area of the reference cell is then given by

$$
\begin{equation*}
A_{r e f 6}=4.5 \cdot \sqrt{3} \cdot W_{2}{ }^{2} \tag{47}
\end{equation*}
$$

The normalized area for hexagonal reference cell, based on the $\mathrm{W} / \mathrm{L}$ of (44) is

$$
\begin{equation*}
A_{r e f, n}=\frac{3}{4} \frac{\sqrt{3} \cdot W_{2}^{2}}{\frac{W_{1}}{d_{1}}+0.4} \tag{48}
\end{equation*}
$$

The gate area and the diffusion area and perimeter are given by

$$
\begin{gather*}
A_{\text {gate } 6}=6 d_{1}\left(\frac{d_{1}}{\sqrt{3}}+d_{2}+\frac{d_{3}}{2}\right)  \tag{49}\\
A_{\text {diff } 6}=\frac{3 \sqrt{3}}{2}\left(W_{1}^{2}+3 \cdot W_{2}^{2}\right)  \tag{50}\\
A_{\text {peri6 }}=6 \cdot\left(W_{1}+W_{2}\right) \tag{51}
\end{gather*}
$$

## Performance comparison

A quantitative comparison of the performance of the reference cells for the alternative layout strategies will be made in this section. The alternating bar structure of Fig. 2 will serve as a reference and area savings of all other structures will be compared with that of the alternating bar structure.

Table 3 shows the area comparison for several different design rule scenarios (defined in Table 2). Column 3 consists of typical design rules in the TSMC process. Considering this scenario, it is seen that a $40 \%$ reduction in area is achievable with the Waffle structure, a $29 \%$ reduction is achievable with the Star Zag structure and nearly a $43 \%$ reduction is achievable with the Fingered-Waffle structure. These substantial reductions in area are achieved while still maintaining a large number of via contacts and a small source resistance.

A comparison between column 1 and column 2 is also of interest. Column 1 shows the results obtained used standard MOSIS design rules and column 2 shows the results obtained using a hypothetical modified MOSIS process. The only change in this modified process is to change the minimum diffusion width from 3 to 2 . As can be seen from table 3 , for almost all the configurations, the modified process outperforms the original process. It thus becomes possible to use the formulas presented earlier as a guide for developing processes that are capable of area-efficient transistors.

A comparison of the diffusion areas of the different layout styles can also be made. Table 4 compares the total diffusion area associated with each reference cell. The numbers shown are normalized by dividing the diffusion area of each reference cell by its respective $W / L$. The results show that significant reduction of total diffusion area over the alternating bars style is possible for waffle, star-zag, modified waffle, and hexagonal styles. The zipper structure, however, has a higher diffusion area associated with it

Table 2: Different scenarios for layout rules

|  | MOSIS | Modified | TSMC |
| :--- | :--- | :--- | :--- |
| d1 | 2 | 2 | 1.7 |
| d2 | 3 | 2 | 2.1 |
| d3 | 2 | 2 | 2.1 |
| d4 | 2 | 2 | 1.7 |
| d5 | 1.5 | 1.5 | 1.0 |
| d6 | 2 | 2 | 2.1 |

Table 3: Area comparison of different layout structures with Alt. Bars structure

|  |  | MOSIS | Modified | TSMC |
| :--- | :--- | ---: | ---: | ---: |
| 1. Alt. Bars |  | 16.0 | 16.0 | 12.5 |
| 2. Waffle |  | 9.8 | 9.8 | 7.5 |
|  | \%incr. | -38.9 | -38.9 | -39.7 |
| 3. Zipper |  |  |  |  |
| Normal |  | 16.7 | 15.5 | 12.3 |
|  | \%incr. | 4.2 | -3.2 | -1.2 |
| Deep |  | 14.3 | 12.5 | 10.2 |
|  | \%incr. | -10.7 | -21.6 | -18.4 |
| Infinite |  | 10.0 | 8.0 | 6.6 |
|  | \%incr. | -37.5 | -50.0 | -47.1 |
| 4. Star Zag |  | 13.3 | 10.8 | 8.9 |
|  | \%incr. | -16.9 | -32.3 | -29.0 |
| 5.Fingered | Waffle |  |  |  |
| $\boldsymbol{x}=\mathbf{2}$ |  | 11.3 | 9.9 | 7.9 |
|  | \%incr. | -29.6 | -37.8 | -36.4 |
| $\boldsymbol{x}=\mathbf{4}$ |  | 10.9 | 9.4 | 7.6 |
|  | \%incr. | -31.7 | -41.4 | -39.3 |
| $\boldsymbol{x}=\mathbf{1 0}$ |  | 10.5 | 8.7 | 7.1 |
|  | \%incr. | -34.3 | -45.4 | -42.8 |
| $\mathbf{6 . ~ H e x a g o n a l ~}$ |  | 21.5 | 19.2 | 15.4 |
|  | \%incr | 34.4 | 20.2 | 23.1 |

Table 4: Normalized diffusion comparison of different layout structures with Alt. Bars structure

|  |  | MOSIS | Modified | TSMC |
| :--- | :--- | ---: | ---: | ---: |
| 1. Alt. Bars |  | 12 | 12 | 9.5 |
| 2. Waffle |  | 5.5 | 5.5 | 4.4 |
|  | \%incr. | -54.2 | -54.2 | -53.9 |
| 3. Zipper |  |  |  |  |
| Normal |  | 11.7 | 10.3 | 8.6 |
|  | \%incr. | -2.8 | -14 | -9.9 |
| Deep |  | 9.6 | 7.8 | 6.8 |
|  | \%incr. | -19.6 | -34.6 | -29.2 |
| 4. Star Zag |  | 8.6 | 6.0 | 5.4 |
|  | \%incr. | -28.3 | -49.9 | -43.7 |
| 5.Fingered | Waffle |  |  |  |
| $\boldsymbol{x}=\mathbf{2}$ |  | 6.2 | 4.8 | 4.2 |
|  | \%incr. | -48.4 | -60.3 | -56.4 |
| $\boldsymbol{x}=\mathbf{4}$ |  | 6.1 | 4.5 | 4.0 |
|  | \%incr. | -48.8 | -62.1 | -57.8 |
| $\boldsymbol{x}=\mathbf{1 0}$ |  | 6.1 | 4.3 | 3.9 |
|  | \%incr. | -49.3 | -64.3 | -59.5 |
| $\mathbf{6 . ~ H e x a g o n a l ~}$ |  | 9.4 | 8.3 | 7.7 |
|  | \%incr | -21.6 | -30.6 | -19.4 |

## Complete transistor

With the reference cells characterized, we can now extend those formulas to obtain (W/L) $)_{\text {eff }}, \mathrm{A}_{\text {ref }}$, and the diffusion area and perimeter for complete transistors built by the interconnection of the reference cells. With the aid of these formulas, it will be easier to compare the area occupied by the transistor of a given size.

For a large transistor, the actual W/L, area occupied, and total diffusion area using a particular layout style can be approximately predicted by the formulas for that style presented in the previous section. The predicted result will have errors for smaller sized transistors due to increased effect of periphery area added for design rules compliance. The choice of shape of each reference cell will determine how much area needs to be added or removed to make the complete transistor free of design rule violations. Detailed derivation of structures described earlier will be presented next.

## Alternating Bar Structure

A complete transistor can be obtained without design rules violations by interconnecting the reference cell of Fig. 2 b with only one change: Poly at the top of the cell must be removed from the top row of the transistor as shown in Fig. 8. The area for a complete transistor in terms of rows of columns of the reference cells is given by

$$
\begin{equation*}
\text { Area }_{\text {total } 1}=A_{\text {ref } 1} \cdot \text { rows } \cdot \text { columns }-\Delta A_{1} \tag{52}
\end{equation*}
$$



Fig. 8: Complete alternating bars
where $\mathrm{A}_{\text {refl }}$ is given by (9) and

$$
\begin{equation*}
\Delta A_{1}=\left[d_{1} \cdot\left(d_{3}+d_{6}\right)\right] \cdot \text { columns } \tag{53}
\end{equation*}
$$

Removing the poly from the top row will also reduce the effective W/L of the complete transistor. The total W/L that takes that into account is given by

$$
\begin{equation*}
\left(\frac{W}{L}\right)_{\text {Lotal } 1}=\left(\frac{W}{L}\right)_{\text {eff } 1} \cdot \text { rows } \cdot \text { columns }-\Delta\left(\frac{W}{L}\right)_{1} \tag{54}
\end{equation*}
$$

where $(W / L)_{\text {effl }}$ is given by (10) and

$$
\begin{equation*}
\Delta\left(\frac{W}{L}\right)_{1}=\left(\frac{d_{3}+d_{6}}{d_{1}}\right) \cdot \text { columns } \tag{55}
\end{equation*}
$$

With the formulas for the effective (W/L) known, it is easy to work backwards to find the needed number of rows and columns to yield the desired (W/L). To make the calculation simpler and to achieve a desired aspect ratio, we can define either rows or columns in terms of the other. For instance, for a nearly square transistor we can define 1 row $=2$ columns. This would yield a simplified (54) with the number of columns as the only unknown.

The total diffusion area for the transistor using alternating bars reference cell can be found in terms of the number of rows and columns of the reference cell

$$
\begin{equation*}
A_{\text {diff toual } 1}=A_{\text {diff } 1} \cdot \text { rows } \cdot \text { columns } \tag{56}
\end{equation*}
$$

Similarly, the total diffusion perimeter can then be given by

$$
\begin{equation*}
P_{\text {diff totad1 }}=P_{\text {diff } 1} \cdot \text { rows } \cdot \text { columns }+\Delta P_{\text {diff } 1} \tag{57}
\end{equation*}
$$

$\Delta P_{\text {diff }}$ accounts for the periphery diffusion perimeter in each row and is given by

$$
\begin{equation*}
\Delta P_{\text {diff } 1}=4 \cdot\left(d_{3}+2 d_{4}\right) \cdot \text { rows } \tag{58}
\end{equation*}
$$

## Waffle

A complete transistor built using interconnecting the waffle structure is shown in figure 9. To make the structure design rules compliant, poly has to be removed from the left side of the reference cells in the leftmost column as well as from the top of the reference cells in the top row. Due to these two modifications, the actual area occupied by the transistor is given by

$$
\begin{equation*}
\text { Area }_{\text {total } 2}=A_{\text {ref } 2} \cdot \text { rows } \cdot \text { columns }-\Delta A_{2 a}-\Delta A_{2 b} \tag{59}
\end{equation*}
$$

where $A_{\text {ref2 }}$ is the reference cell area obtained earlier in (15). The correction due to removing poly from the leftmost column is given by

$$
\begin{equation*}
\Delta A_{2 a}=2\left(d_{1}+d_{3}+2 d_{4}\right) \cdot d_{1} \cdot \text { rows } \tag{60}
\end{equation*}
$$

Similarly, the correction in the area calculation due to poly removal from the top row is given by

$$
\begin{equation*}
\Delta A_{2 b}=d_{1}\left(d_{1}+d_{3}+2 d_{4}\right) \cdot(\text { columns }-1)+d_{1} \cdot\left(d_{3}+2 d_{4}\right) \tag{61}
\end{equation*}
$$

The removal of poly also affects the total (W/L) $)_{\text {eff }}$ of the complete transistor. The actual (W/L) is given by

$$
\begin{equation*}
\left(\frac{W}{L}\right)_{\text {toul } 2}=\left(\frac{W}{L}\right)_{e f f 2} \cdot \text { rows } \cdot \text { columns }-\Delta\left(\frac{W}{L}\right)_{2 a}-\Delta\left(\frac{W}{L}\right)_{2 b} \tag{62}
\end{equation*}
$$

The correction due to removing poly from the left side is given by

$$
\begin{equation*}
\Delta\left(\frac{W}{L}\right)_{2 a}=2 \cdot\left(\frac{d_{3}+2 d_{4}+0.55 \cdot d_{1}}{d_{1}}\right) \cdot \text { rows } \tag{63}
\end{equation*}
$$



Fig.9: Complete waffle
The change in the (W/L)eff due to the correction in the top row is

$$
\begin{equation*}
\Delta\left(\frac{W}{L}\right)_{2 b}=\left(\frac{d_{3}+2 d_{4}+0.55 \cdot d_{1}}{d_{1}}\right) \cdot(\text { columns }-1)+\frac{d_{3}+2 d_{4}}{d_{1}} \tag{64}
\end{equation*}
$$

We need to know the total diffusion area and perimeter to estimate the parasitic capacitances associated with the complete transistor. For the waffle structure, they are given by (in terms of their reference cell diffusion and perimeter)

$$
\begin{align*}
& A_{\text {diff. toout } 2}=A_{\text {diff } 2} \cdot \text { rows } \cdot \text { columns }  \tag{65}\\
& P_{\text {diff.total } 2}=P_{\text {diff } 2} \cdot \text { rows } \cdot \text { columns } \tag{66}
\end{align*}
$$

As a good approximation, half of the total area and perimeter calculated from (65) and (66) can be assigned to drain and source each.

## Zipper

To build a complete transistor using the zipper reference cells, we need to make two additions to the stack of the reference cells. First, we need to add enough diffusion at the bottom to form either the drain or the source contacts. Second, we need to add diffusion regions to the ends of each row for design rules compliance. These two additions are shown in Fig. 10. The total area is then given by

$$
\begin{equation*}
\text { Area }_{\text {toual } 3}=A_{\text {ref } 3} \cdot \text { rows } \cdot \text { columns }+\Delta A_{3 a}+\Delta A_{3 b} \tag{67}
\end{equation*}
$$

The correction due to adding the extra diffusion and contacts to the bottom row is given by


Fig.10: Complete zipper

$$
\begin{equation*}
\Delta A_{3 a}=\left[\left(d_{3}+d_{4}+d_{5}\right)\left(2 d_{1}+2 d_{2}\right)\right] \cdot \text { columns } \tag{68}
\end{equation*}
$$

Adding diffusion to both ends of each row results in the change given by

$$
\begin{equation*}
\Delta A_{3 b}=\left((x+3) d_{1}+d_{3}+2 d_{4}\right) \cdot d_{2} \cdot \text { rows } \tag{69}
\end{equation*}
$$

For the zipper structure, the change in (W/L)eff due to the needed additions is very small. The only change is the small extension of poly needed on both ends of each row. The total (W/L) and the correction are given by (70) and (71) below.

$$
\begin{align*}
\left(\frac{W}{L}\right)_{\text {total3 }}= & \left(\frac{W}{L}\right)_{\text {eff } 3} \cdot \text { rows } \cdot \text { columns }+\Delta\left(\frac{W}{L}\right)_{3}  \tag{70}\\
& \Delta\left(\frac{W}{L}\right)_{3}=\left(\frac{d_{2}}{d_{1}}\right) \cdot \text { rows } \tag{71}
\end{align*}
$$

The total diffusion area associated with the complete transistor can also be found. The total diffusion area with the corrections due to diffusion addition to the bottom and each row are given by (72), (73), and (74) respectively

$$
\begin{gather*}
A_{\text {diff } . \text { total } 3}=A_{\text {diff } 3} \cdot \text { rows } \cdot \text { columns }+\Delta A_{\text {diff } 3 a}+\Delta A_{\text {diff } 3 b}  \tag{72}\\
\Delta A_{\text {diff } 3 a}=\left(d_{3}+d_{4}+d_{5}\right)\left(2 d_{1}+2 d_{2}\right) \cdot \text { columns }  \tag{73}\\
\Delta A_{\text {diff } 3 b}=2\left((x+2) d_{1}+d_{3}+2 d_{4}\right) \cdot d_{2} \cdot \text { rows } \tag{74}
\end{gather*}
$$

The total diffusion parameter is given by

$$
\begin{equation*}
P_{\text {diff toolal3 } 3}=P_{\text {diff } 3} \cdot \text { rows } \cdot \text { columns }+\Delta P_{\text {diff } 3 a}+\Delta P_{\text {diff } 3 h} \tag{75}
\end{equation*}
$$

$\Delta P_{d i f j a}$ and $\Delta P_{d i f 3 b b}$ account for the periphery diffusion perimeter in each row and column and are given by

$$
\begin{gather*}
\Delta P_{\text {diff } 3 a}=\left(4 d_{1}+3 d_{2}\right) \cdot \text { columns }+2\left(d_{2}+d_{3}+2 d_{4}\right)  \tag{76}\\
\Delta P_{\text {diff } 3 b}=2\left((x+2) d_{1}+d_{3}+2 d_{4}\right) \cdot \text { rows } \tag{77}
\end{gather*}
$$

For a large transistor, half of the total diffusion area can be allocated to the drain and source each.

## Star Zag

A complete transistor can be made by joining the star zag reference cells together. A transistor with rows=columns $=2$ is shown in Fig. 11. Notice the two changes in the reference cells in the bottom row and right column to make the transistor design rule compliant. Specifically, a strip with a width equal to that of a minimum poly width is removed from the bottom row cells and a wider portion is sliced from the right side of the right column cells. These changes result in the total area given by

$$
\begin{equation*}
\text { Area }_{\text {toolal } 4}=A_{\text {ref } 4} \cdot \text { rows } \cdot \text { columns }-\Delta A_{4 a}-\Delta A_{4 b} \tag{78}
\end{equation*}
$$



Fig.11: Complete star zag
The correction due to area taken from the bottom row cells is given by

$$
\begin{equation*}
\Delta A_{4 a}=2\left(4 d_{1}+2 d_{2}\right) \cdot d_{1} \cdot(\text { columns }-1)+\left(5 d_{1}+4 d_{2}\right) \cdot d_{1} \tag{79}
\end{equation*}
$$

The correction due to removing of a portion from the right column cells is given by

$$
\begin{equation*}
\Delta A_{4 b}=3 d_{1}\left(3 d_{1}+3 d_{2}\right) \cdot \text { rows } \tag{80}
\end{equation*}
$$

Since the area removed earlier contains poly regions, the (W/L) eff of the transistor is reduced. The corrected total (W/L) $)_{e f f}$ is given by

$$
\begin{equation*}
\left(\frac{W}{L}\right)_{\text {total } 4}=\left(\frac{W}{L}\right)_{\text {eff } 4} \cdot \text { rows } \cdot \text { columns }-\Delta\left(\frac{W}{L}\right)_{4 a}-\Delta\left(\frac{W}{L}\right)_{4 b} \tag{81}
\end{equation*}
$$

The correction due to the bottom row cells' modification is given by

$$
\begin{equation*}
\Delta\left(\frac{W}{L}\right)_{4 a}=\left(\frac{6.2 d_{1}+2 d_{2}}{d_{1}}\right) \cdot(\text { columns }-1)+\left(\frac{3.65 d_{1}+2 d_{2}}{d_{1}}\right) \tag{82}
\end{equation*}
$$

The reduction in (W/L) eff due to slicing part of the right column cells is given by

$$
\begin{equation*}
\Delta\left(\frac{W}{L}\right)_{4 b}=\left(\frac{7.2 d_{1}+2 d_{2}}{d_{1}}\right) \cdot \mathrm{rows} \tag{83}
\end{equation*}
$$

By expressing either the rows or the columns in term of the other in (81), the required number of rows and columns for a desired (W/L) can be found.

To find the diffusion area for the complete transistor, we need to adjust the total area to account for the smaller reference cells in the rightmost column and the bottom row. The total diffusion area and the above mentioned adjustments are given by

$$
\begin{gather*}
A_{\text {diff fotal } 4}=A_{\text {diff } 4} \cdot \text { rows } \cdot \text { columns }-\Delta A_{\text {diff } 4 a}-\Delta A_{\text {diff } 4 b}  \tag{84}\\
\Delta P_{\text {diff } 4 a}=4 d_{1} \cdot \text { columns }  \tag{85}\\
\Delta P_{\text {diff } 4 b}=\left(14 d_{1}-2 d_{2}\right) \cdot \text { rows } \tag{86}
\end{gather*}
$$

Similarly, the total diffusion perimeter for a complete transistor for the Star Zag case is given by

$$
\begin{gather*}
P_{\text {diff foral } 4}=P_{\text {diff } 4} \cdot \text { rows } \cdot \text { columns }-\Delta P_{\text {diff } 4 a}-\Delta P_{\text {diff } 4 b}  \tag{87}\\
\Delta P_{\text {diff } 4 a}=4 d_{1} \cdot \text { columns }  \tag{88}\\
\Delta P_{\text {diff } 4 b}=\left(14 d_{1}-2 d_{2}\right) \cdot \text { rows } \tag{89}
\end{gather*}
$$

For a large transistor, drain and source can each be assigned half of the total diffusion area for parasitic capacitance calculations.

## Fingered Waffle

A complete transistor obtained by the interconnection of fingered waffle reference cells ( 3 rows and 2 columns) is shown in Fig. 12. In a manner similar to Star Zag configuration discussed earlier, two changes need to be made to make the transistor design rule compliant. First, a strip equal to minimum poly width is removed from the bottom row. Second, a larger portion is removed from each row. The total area for this configuration is given by

$$
\begin{equation*}
\text { Area }_{\text {total } 5}=A_{\text {ref } 5} \cdot \text { rows } \cdot \text { columns }-\Delta A_{5 a}-\Delta A_{5 b} \tag{90}
\end{equation*}
$$

The area correction due to change in the bottom row is given by

$$
\begin{equation*}
\Delta A_{5 a}=2\left((x+1) d_{1}+d_{3}+2 d_{4}\right) \cdot d_{1} \cdot(\text { columns }-1)+\left((x+1) d_{1}+2 d_{3}+4 d_{4}\right) \cdot d_{1} \tag{91}
\end{equation*}
$$

The correction factor to account for the change of area in each row is given by

$$
\begin{equation*}
\Delta A_{5 b}=(x+1) d_{1} \cdot\left(2 d_{1}+2 d_{2}\right) \cdot \text { rows } \tag{92}
\end{equation*}
$$

Since the complete transistor uses some reference cells that are smaller in size, the actual (W/L) eff has to be adjusted. The total (W/L) and the needed adjustments are given by

$$
\begin{equation*}
\left(\frac{W}{L}\right)_{\text {total } 5}=\left(\frac{W}{L}\right)_{e f f 5} \cdot \operatorname{rows} \cdot \operatorname{columns}-\Delta\left(\frac{W}{L}\right)_{5 a}-\Delta\left(\frac{W}{L}\right)_{5 b} \tag{93}
\end{equation*}
$$

The change in W/L due to removing poly from the bottom row is

$$
\begin{equation*}
\Delta\left(\frac{W}{L}\right)_{5 a}=\left(\frac{2 x \cdot d_{1}+2 d_{3}+4 d_{4}+2 \cdot 0.55 d_{1}}{d_{1}}\right) \cdot(\text { columns }-1)+\left(\frac{x \cdot d_{1}+2 d_{3}+4 d_{4}+0.55 d_{1}}{d_{1}}\right) \tag{94}
\end{equation*}
$$



Fig.12: Complete modified waffle

The correction applied due to adjustment to the end of each row is

$$
\begin{equation*}
\Delta\left(\frac{W}{L}\right)_{5 b}=\left(\frac{(2 x-1) d_{1}+2 d_{2}+3 \cdot 0.55 \cdot d_{1}}{d_{1}}\right) \cdot \text { rows } \tag{95}
\end{equation*}
$$

The total diffusion area is given by

$$
\begin{gather*}
A_{\text {diff total } 5}=A_{\text {diff } 5} \cdot \text { rows } \cdot \text { columns }-\Delta A_{\text {diff } 5}  \tag{96}\\
\Delta A_{\text {diff } 5}=\left(2 x \cdot d_{1} \cdot d_{2}\right) \cdot \text { rows } \tag{97}
\end{gather*}
$$

Similarly, the diffusion perimeter is given by

$$
\begin{gather*}
P_{\text {diff totat } 5}=P_{\text {diff } 5} \cdot \text { rows } \cdot \text { columns }-\Delta P_{\text {diff } 5}  \tag{98}\\
\Delta P_{\text {diff } 5}=4 x \cdot d_{1} \cdot \text { rows } \tag{99}
\end{gather*}
$$

As an approximation for large transistors, half of the diffusion source and perimeter may be assigned to source and drain each.

## Hexagonal

A complete transistor using the hexagonal reference cell can be built to have a rectangular or a non-rectangular aspect ratio. The latter will be assumed here. It should also be noted that the reference cell used here is not optimized and better area utilization may be possible. An example of rectangular aspect ratio transistor composed of individual reference cells is shown in Figure 7. The area and effective $\mathrm{W} / \mathrm{L}$ of complete transistor are given by

$$
\begin{gather*}
\text { Area }_{\text {totat } 6}=A_{\text {ref } 6} \cdot\left(\text { rows } \cdot \text { columns }+\frac{2}{3}(\text { rows }+ \text { columns })\right)  \tag{100}\\
 \tag{101}\\
\left(\frac{W}{L}\right)_{\text {total } 6}=\left(\frac{W}{L}\right)_{\text {eff } 6} \cdot \text { rows } \cdot \text { columns }
\end{gather*}
$$

For a given number of rows and columns, the total diffusion area and perimeter are given by

$$
\begin{align*}
A_{\text {diff ftoal } 6} & =3 \sqrt{3}\left(\text { rows } \cdot \text { columns } \cdot\left(\frac{W_{1}^{2}}{2}+W_{2}^{2}\right)+(\text { rows }+ \text { columns }) \cdot W_{2}^{2}\right)  \tag{102}\\
P_{\text {diff t.tual } 6} & =6 \cdot \text { rows } \cdot \text { columns } \cdot\left(W_{1}+W_{2}\right)+(8 \cdot(\text { rows }+ \text { columns })+2) \cdot W_{2} \tag{103}
\end{align*}
$$

For a large transistor, source area is approximately three times larger than the drain area.

## Conclusions

MOS transistors with very large effective W/L are needed in applications such as switches. Alternate layout structures were presented that have the potential of achieving a higher effective W/L ratio for the same area utilized by traditional layout structures. Traditional non-rectangular transistor structures were analyzed to obtain effective W/L. Alternate layout structures were analyzed for comparison with traditional structures. It was shown that substantial reduction in area is achievable by using Waffle structures or modified Waffle structures. A reduction in area of over $40 \%$ and a reduction in diffusion area of more than $50 \%$ were demonstrated for a typical process. Although these structures are geometrically intricate, closed-form design equations were presented that can facilitate the utilization of these structures.

## Acknowledgments

Support for this project was provided, in part, by Texas Instruments Inc. and by Rockwell International.

## References

[1] P. Grignoux and R. L. Geiger, "Modeling of MOS transistors with nonrectangular-gate geometries," IEEE Trans. on Electron Devices, Vol 29, pp.1261-1269, August 1982.
[2] K. Laker and W. Sansen, Design of Analog Integrated Circuits and Systems. New York, NY: McGraw Hill, 1994
[3] D.A. Grant and J.G. Gowar, Power MOSFETS Theory and Applications. New York, NY: Wiley, 1989.
[4] J. Bastos, M. Steyaert, B. Graindourze, W. Sansen, "Matching of MOS transistors with different layout styles", Proc. IEEE Int. Conference on Microelectronic Test Structures, vol. 9, pp17-18, 1996.
[5] S. R. Vemuru, "Layout comparison of MOSFETs with large W/L ratios", Electron Letters, Vol. 28, pp.2327-2329, December 1992.
[6] S. Q. Malik and R. L. Geiger, "Minimization of area in low-resistance MOS switches," Proceedings of the 43rd IEEE Midwest Symposium on Circuits and Systems, August 2000
[7] A. Van den Bosch, M. Steyaert, and W. Sansen, "A high density, matched hexagonal transistor structure in standard CMOS technology for high-speed applications," IEEE Trans. on Semiconductor Manufacturing, Vol 13, no. 2, pp.167-172, May 2000
[8] L. Baker, et. al., "A 'waffle` layout technique strengthens the ESD hardness of the NMOS output transistors," 1989 EOS/ESD Symposium Proceedings, pp.175-181

## CHAPTER 6. CONCLUSION

The papers presented in this dissertation describe techniques for switched capacitor circuits and for reducing the area of extreme ratio transistors. A brief summary of each work and the contribution is provided next.

## Capacitor sharing and scaling technique for reduced power in pipelined <br> ADCs


#### Abstract

Pipelined ADCs are used for medium to high speed applications with resolutions higher that 8 bits. The MDAC forms an integral form of a single stage of a pipeline ADC and is typically implemented using switched capacitor techniques. In such implementations, the opamp is a major source of power dissipation. Any power savings achieved in the opamp of a pipeline ADC, especially in the first few stages, can significantly reduce the total power dissipation of the circuit. The proposed technique is based on the observation that the residue voltage stored across the feedback capacitor at the end of an amplify phase in an MDAC is typically not utilized. Instead, the next stage's capacitors are used to sample the same residue voltage that is amplified later. Additionally, it is known that power savings in a pipelined $A D C$ are possible if capacitors in subsequent stages are scaled by the interstage gain. The proposed technique eliminates the sampling capacitors of the second of any two consecutive stages and reuses the charge stored on the feedback capacitor of the first stage. That feedback capacitor can be replaced with a network of capacitors and the residue voltage across this compound capacitor is reused in the second stage. The network automatically scales the capacitors by the interstage gain while moving it to the second stage. The technique can also be modified to share opamps between two stages for additional power savings.


## Contributions

The contributions for this project are:

- An architecture level technique that allows reuse of charge on a capacitor by sharing a capacitor between two consecutive stages. The technique is not limited by the power supply of the ADC and can be used to modify an existing ADC to achieve lower power.
- The proposed technique can reduce the load of the opamp by more than $40 \%$ of the conventional case. This reduction can be used to either cut down on the capacitor sizes resulting in reduced
power or to achieve a faster speed of operation if the capacitor sizes cannot be made smaller due to noise or matching constraints.


## A capacitor sharing technique for RSD cyclic ADC

Cyclic ADCs are used in applications that require low power dissipation and small area while achieving medium levels of accuracies. Due to their nature, achievable speeds are typically lower than those achievable by pipelined ADC. However, the traditional cyclic ADC implementations are very similar to the first two stages of a pipeline ADC. This resemblance allows the extending of the technique proposed for the pipeline ADC earlier to cyclic ADCs . In the proposed technique, the residue charge stored across a set of capacitors in the first cycle is reused in the next clock cycle. The technique can be combined with other area and power saving techniques such as opamp sharing. Error correction techniques that can tolerate large offset errors in the comparators are combined with the implementation. A 10 -bit, 2.3 MHz cyclic ADC in AMI $0.5 \mu$ process was implemented that used shared opamp as well as error correction technique. Simulation results show a THD of -76.11 dB and an SFDR of 74.95 dB .

## Contributions

The contributions in this work are:

- The technique proposed for the pipeline ADC was extended for use with cyclic ADCs. The charge stored on a set of feedback capacitors in the first cycle is reused during the second cycle. Similar to the pipeline case, the technique allows making architecture level changes independent of the supply voltage values.
- It was shown that the conventional circuit could be modified to achieve either a smaller area and reduced power by reducing the capacitor sizes or faster operation by reducing the loading on the opamp. To a first order, $50 \%$ reduction in dynamic power dissipation is possible.


## A low temperature sensitivity switched-capacitor current reference

Voltage and current references are used in virtually all analog integrated circuits. It is desirable that these reference quantities remain as constant with temperature variations as possible. Unlike voltage references that can be derived from intrinsic physical values of the process, no intrinsic current
reference is available in CMOS. However, a reference current that has low temperature sensitivity can be obtained using switched capacitor technique. A highly stable clock is usually available to systems that consist of switched capacitor circuits. The proposed technique utilizes this stable clock in conjunction with linear capacitors available in many processes to generate a current that has a very low sensitivity to temperature variations. The technique involves transferring a fixed amount of charge onto a circuit node whose time average value is kept constant using a feedback network. Simulated results presented in the paper showed the variation in the reference current to be less than $0.029 \%$ over a temperature range of $-40^{\circ} \mathrm{C}$ to $125^{\circ} \mathrm{C}$.

## Contributions

The contributions of this work are:

- A technique of generating a reference current with very low temperature sensitivity. The proposed switched capacitor technique uses a stable external clock and linear capacitors on chip to generate a reference current with low temperature sensitivity.
- Practical design issues were anticipated and possible solutions were presented. Techniques to increase the output resistance, improving the settling time, and reducing ripple in the reference current were demonstrated.


## Area efficient layout strategies for extreme-ratio MOS transistors

MOS transistors are often used as switches and in pad drivers that drive large external loads. These applications typically require the effective resistance of the transistor to be very low. The effective resistance of a MOS transistor operated as a switch is affected by several parameters. The four that generally receive the most attention are the W/L ratio, the excess bias, the series diffusion resistance, and the contact resistance. For applications requiring low resistance, large effective W/L ratios are used along with multiple contacts to the drain and source diffusions. Instead of using one huge poly to form the gate of the transistor, the large effective W/L ratio is traditionally achieved by connecting multiple transistors of smaller W/L ratio in parallel. However, these transistors are not the most area efficient, as shown in the paper. Instead of using multiple fingers, alternate structures such as the waffle and its variants result in higher area efficiency for a given W/L.

Parasitic capacitances associated with the diffusion regions of transistors can become a limiting factor in the high speed performance of the design. For such designs, any reduction in these parasitic
capacitances is highly desirable. Using alternate layout structures results in reduced diffusion area and perimeters, i.e., smaller parasitic capacitance, enabling higher speeds of operation.

## Contributions

The contributions of this work are:

- Alternate existing and layout structures for MOS transistor with extremely large effective $\mathrm{W} / \mathrm{L}$ ratios were presented. Closed form expressions based on the physical design rules were derived. The expressions can be used to determine the number of building blocks needed to create a complete transistor.
- Expanded expressions for the complete transistor were presented. These expressions for the $(\mathrm{W} / \mathrm{L})_{\text {effective }}$ and total area included periphery effects of putting together a complete large transistor.
- The derived expressions were used to compare the performance of different layout structures and compared with the traditional structure. The equations presented in the paper used design rules as variables. As a result, it becomes possible to compare the different layouts conforming to different sets of design rules. Using design rule based expressions also makes it possible to evaluate a new process for area efficiency of very large transistors.


[^0]:    ${ }^{1}$ Held in August 2001 at Helsinki University of Technology (HUT), Espoo, Finland.

