平成２７年度　博士論文

Low Power Circuit Design Methodology with Layout Dependent Effect-aware Leakage Current Reduction

(レイアウト依存効果を考慮したリーク電流削減を伴う低消費電力設計手法)

指導教員　中武　繁寿　教授

北九州市市立大学大学院国際環境工学研究科

情報工学専攻　コンピュータシステムコース

2010DCB401　　陳　功

Summary

As semiconductor manufacturing processing scaling down, leakage current of CMOS circuits is becoming a dominant contributor to power dissipation. Based on analyzing of low-power issues, we conclude that the leakage current, especially the STI induced junction leakage dominates the power consumption below 65nm process, especially in low frequency circuitries.

This paper provides an efficient leakage current reduction (LCR) technique for low-power and low-frequency circuit designs in terms of design rules and layout parameters related to layout dependent effects. We analyzed STI induced junction current model. Based on the analysis, we concluded and proposed LDE-aware LCR design procedure. The investigation of single transistor I-V characteristics and simulation of Inverter power consumption support our point of view.

We address the LCR technique both for analog and digital circuits, and present a design case when applying the LCR technique to a successive-approximation-register (SAR) analog-to-digital converter (ADC), which typically employs analog and digital transistors. In the practical example of SAR ADC, we discuss the design procedure step by step and implement LDE-aware LCR in fundamental blocks. Moreover, we introduce the basic design specifications and methodologies for the fundamental blocks of SAR ADCs. There are a comparator energy model and a liner bootstrapped switch among them from our previous work. Some state-of-the-art low power design technologies also apply in this 11-bit SAR ADC design. In the post-layout simulation results by HSPICE, an SAR-ADC with the LCR technique achieves 38.6-nW as the total power

consumption. Comparing with the design without the LCR technique, we attain about 30% total energy reduction.

内容の要旨

半導体微細加工技術の進歩に伴い、CMOS 回路の消費電力におけるリーク電流の割合が無視できなくなってきている。本研究では、低電力技術に係る調査に基づき、65nm 製造プロセス以降では、特に、低周波回路に対して、STI がデバイス接合部に及ぼすリーク電流がより支配的となっている事実を明らかにした。さらに、低電力・低周波回路設計に対して、我々は、レイアウト依存効果に関連する設計ルールと設計パラメータに換算される効果的なリーク電流低減手法を提供する。

まず我々は、STI が接合部電流に及ぼす影響のモデルを解析し、レイアウト依存効果を考慮したリーク電流低減のために、設計上行うべきことを明らかにした。単体のトランジスタの電流・電圧特性と CMOS インバータの電力消費シミュレーションにおける考察は、我々の考えが妥当性を示唆している。

次に我々は、アナログとディジタル回路の両方に対する消費電力低減技術を述べ、提案技術を典型的な逐次比較型 ADC 回路(SAR-ADC)に適用事例を示す。SAR-ADC の実例では、提案するレイアウト依存効果を考慮したリーク電流低減手法を、基本ブロックレベルごとに、設計上の流れにあわせて各手続きを議論している。

さらに、コンパレータのエネルギーモデルと線形ブートストラップスイッチに対する考察から、SAR-ADC の基本ブロックに対して、基準とすべき設計仕様とそれに対する設計方式を導入する。HSPICE を用いたポストレイアウトシミュレーションでは、提案するリーク電流低減手法を適用した SAR-ADC は、全消費電力として 38.6nW を達成し、提案手法を適用しない場合と比較して、約 30%の電力削減が実現できることを示した。

Acknowledgments

Completion of this doctoral dissertation was possible with the support of several people. I wish to express my gratitude to my advisor, Dr. Shigetoshi NAKATAKE for guiding me in my work and backing me philosophically and financially. He is an exceptional human being who serves as my inspiration for a career in academics.

I am extremely grateful for the extensive technical and moral support from Dr. Bo YANG and Dr. Jing LI. They have been gave me a lot of invaluable help in various tapeouts. I would like to thank Dr Qing DONG and Dr Mingyu LI for being members on my reading committee and spending valuable time outside class in evaluating my thesis work. I would like to thank Dr. Yu ZHANG for his extensive help in digital design.

Finally, I would have never made it to this point without the constant support of my parents and parents-in-law. Most importantly I would like to thank my wife, Dr. Qianyu DONG, She has been an integral part of my life and has made me stronger. I am truly blessed to have her in my life. Their constant support and motivation is something I will always admire.

Table of contents

Chapter -1 Preface 3

1.1 Background 3

1.2 Motivation 5

1.2.1 Power Consumption in CMOS Circuits 5

1.2.2 Analyzed Leakage Current 8

1.3 Dissertation Organization and Contributions 13

References 16

Chapter -2 LDE-aware Leakage Current Reduction Technique 19

2.1 Preliminary Investigation 19

2.2 STI Specifications 20

2.3 STI Induced Junction Current Model 24

2.4 Descriptions of Technique of LDE-aware LCR 27

References 33

Chapter -3 Nano-watt CMOS Reference Circuit Retargeting Methodology with LDE-aware LCR 35

3.1 Preliminaries 35

3.2 A conventional method to derive reference current and voltage based on re-defined Vth 38

3.3 Analyzing SCM and SBC based on ACM model 43

3.4 Retargeting flow with tuning procedure 49

3.5 Leakage current compensation by start-up circuit 51

3.6 LDE-aware LCR be utilized in the reference circuit 52

3.7 Temperature Compensation 53

3.8 Results 55

3.9 Conclusion Remarks 65

References 67

Chapter -4 Low-power and Low-speed SAR ADC Design Methodology with LDE-aware LCR 69

4.1 Design Trend of Analog-to-Digital Converters and Design Gist 69

4.2 SAR ADC Structures and Design Consideration 77

4.3 SAR ADC Architecture Design with Considering LDE-aware LCR 82

4.3.1 Description of Proposed Overall Design Flows 82

4.3.2 Description of SAR ADC Architecture 88

4.4 Implementation of LDE-aware LCR in switch design 92

4.4.1 Top plate sampling (TPS) switch 93
4.4.2 Bottom plate sampling (BPS) switch 94
4.4.3 A Linearity Bootstrapped Switch with Dynamic Bulk Biasing Design 95
4.4.4 LDE-aware LCR Effect on CMOS Switch Design 105
4.5 Implementation of LDE-aware LCR in comparator design 109
4.5.1 Design Basics 110
4.5.2 Proposed comparator energy model 112
4.5.3 LDE-aware LCR effect on comparator design 115
4.6 Implementation of LDE-aware LCR in SAR control logic 117
4.6.1 Synchronous and Asynchronous SAR controller 117
4.6.2 Proposed SAR control logic with LDE-aware LCR 119
4.7 Capacitor array considerations 122
4.8 Results 123
4.8.1 Measured Power Consumption Results 130
4.8.2 Post-Layout Simulation Results of SAR ADC Performance 130
References 132
Chapter -5 Conclusion 136
Publications 138

Chapter -1 Preface

1.1 Background

There has been much interest in recent years in low power circuitry design stemming from the demands of long battery life in portable systems and heat removal in larger, non-portable ones. Low power circuitries are gaining considerable interest from scientific community and more recently the market [1]. The exploding market of portable electronic appliances fuels the demand for complex integrated systems that can be powered by lightweight batteries with long times between re-charges (for instance, the plot in Fig. 1.1 shows the evolution of the worldwide market for mobile phones [2]). Additionally, system cost must be extremely low to achieve high market penetration. Both battery lifetime and system cost are heavily impacted by power dissipation. For these reasons, the last ten years have witnessed a soaring interest in low power design [3], [4].

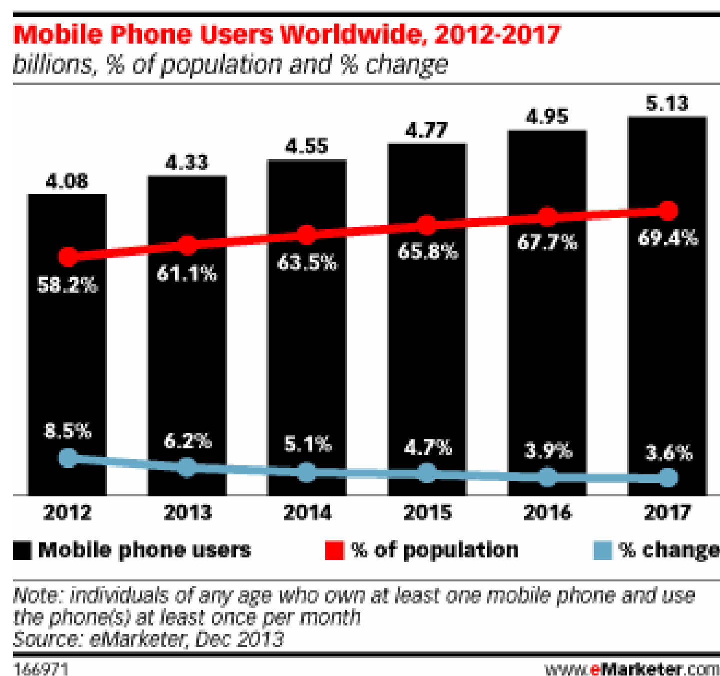


Fig.1.1 Global market for mobile phones

Power dissipation is defined as the rate of energy delivered from the source to the system / device [5]. In battery operated systems, the amount of energy stored within the battery is limited. Therefore, power dissipation is important for portable systems, as it defines the average lifetime of the battery. Unfortunately, battery technology is not expected to improve the battery storage capacity by more than 30% every five years [6]. This is not sufficient to handle the increasing power requirements of portable systems. Low-power devices are expected to have smaller battery size, less weight, and longer battery lifetime.

Power dissipation is also crucial for Deep Sub Micron (DSM) technologies. Advances in CMOS fabrication technology double the number of transistors per chip every two years and double the operating frequency every three years. Consequently, the power dissipation per unit area grows, increasing the chip temperature. This excessive temperature reduces the reliability and lifetime of the circuit. Hence, large cooling devices and expensive packaging are required to dissipate the extra heat. Also, systems with high power dissipation require a special Printed Circuit Board (PCB) technology to deliver large currents from the power supply to the various devices in the system For example, a Pentium III® processor requires an 18A power supply [7], which cannot be handled by a conventional PCB process. Therefore, the price and area of the digital system increase and the transistor density decreases obliterating the advantages of smaller transistor sizes [8]. Furthermore, large current densities cause serious problems. Electro-migration caused by large currents flowing through narrow wires, may produce gaps or bridges in the power rails of the chip with a subsequent permanent damage to the system.

Another reason for low-power design arises from modern automated offices. In 1993, the American Council for Energy Efficient Economy reported that office equipment accounted for 5% of the total commercial energy consumed [9]， and drivers of energy use in building was 40% of total U.S. energy use by the year 2005 [10]. The generation of this energy costs two billion dollars annually and generates pollution equivalent to five million cars. This problem resulted in the development of the "Green Computers" concept to reduce the amount of energy consumed by office equipment.

Though power dissipation is important for portable devices, performance (speed) continues to be the main target for digital designers. Consumers expect higher speed, more functionality, and higher levels of integration, from their cellular phones and hand-holds. To emphasis the importance of speed, researchers use Energy Delay Product (EDP) as an evaluation figure for digital systems. Consequently, reducing power dissipation should not come at the expense of performance. In the meantime, increasing performance while keeping power dissipation constant is also considered to be a low-power design problem.

1.2 Motivation

1.2.1 Power Consumption in CMOS Circuits

CMOS is, by far, the most common technology used for manufacturing ICs. The power consumed by a CMOS circuit is defined as

$$\mathbf{P(t) = I(t)V(t).} \qquad \text{(Eq.1.1)}$$

Where **I(t)** is the instantaneous current provided by the power supply, and **V(t)** is the instantaneous supply voltage. The latter impacts battery lifetime

and heat dissipation system cost, the former constrains power grid and power supply circuits design.

There are three major sources of power dissipation in a CMOS circuit [11]:

$$\mathbf{P(t)} = \mathbf{P_{Switching}(t)} + \mathbf{P_{Short-Circuit}(t)} + \mathbf{P_{Leakage}(t)}. \quad \text{(Eq.1.2)}$$

$\mathbf{P_{Switching}(t)}$ called switching power, is due to charging and discharging capacitors driven by the circuit. $\mathbf{P_{Short-Circuit}(t)}$ called short-circuit power, is caused by the short circuit currents that arise when pairs of PMOS/NMOS transistors are conducting simultaneously. Finally, $\mathbf{P_{Leakage}(t)}$ called leakage power, originates from substrate injection and sub-threshold effects. For older technologies (0.8μm and above), $\boldsymbol{P_{Switching}}\mathbf{(t)}$ was predominant. For deep-submicron processes, $\boldsymbol{P_{Leakage}}\mathbf{(t)}$ becomes more important.

Design for low-power implies the ability to reduce all three components of power consumption in CMOS circuits during the development of a low power electronic product. Optimizations can be achieved by facing the power problem from different perspectives: design and technology. Enhanced design capabilities mostly impact switching and short-circuit power; technology improvements, on the other hand, contribute to reductions of all three components.

Switching power for a CMOS gate working in a synchronous environment is modeled by the following equation:

$$\mathbf{P_{Switching}} = \frac{1}{2} \boldsymbol{C_L V_{DD}^2 f_{Clock} E_{SW}}. \quad \text{(Eq.1.3)}$$

Where $\boldsymbol{C_L}$ is the output load of the gate, $\mathbf{V_{DD}}$ is the supply voltage, $\mathbf{f_{Clock}}$ is the clock frequency and $\mathbf{E_{SW}}$ is the *switching activity* of the gate, defined as the probability of the gate's output to make a logic transition during one

clock cycle. Reductions of $P_{Switching}$ are achievable by combining minimization of the parameters in the equation above.

Glue logic and controllers have much more irregular structure than arithmetic units, and their gate-level implementations are characterized by a wide distribution of path delays. These circuits can be optimized for power by *resizing*. Resizing focuses on fast combinational paths. Gates on fast paths are down-sized, thereby decreasing their input capacitances, while at the same time slowing down signal propagation. By slowing down fast paths, propagation delays are equalized, and power is reduced by joint spurious switching and capacitance reduction. Resizing does not always imply down-sizing. Power can be reduced also by enlarging (or buffering) heavily loaded gates, to increase their output slew rates. Fast transitions minimize short-circuit power of the gates in the fan-out of the gate which has been sized up, but its input capacitance is increased. In most cases, resizing is a complex optimization problem involving a tradeoff between output switching power and internal short-circuit power on several gates at the same time.

Leakage power is already a major concern in current technologies, because it impacts battery lifetime even if the circuit is completely idle. Quiescent power specifications tend to be very tight. In fact, CMOS technology has traditionally been extremely power-efficient when transistors are not switching, and system designers expect low leakage from CMOS chips. To meet leakage power constraints, multiple-threshold and variable threshold circuits have been proposed [12]. In multiple-threshold CMOS, the process provides two different thresholds. Low-threshold transistors are fast and leaky, and they are employed on speed-critical sub-circuits. High-threshold transistors are slower but exhibit low sub-threshold leakage, and they are employed in noncritical units/paths of the chip.

Intel's StrongARM family is introduced in [13]. This design is representative of the class of low-power integrated core and peripherals for personal digital assistants and palmtop computers. Leakage power was the main concern in the StrongARM design, which could not benefit from multiple threshold or variable threshold control. To reduce leakage power in gates outside the critical path, non-minimum length transistors were employed in low-leakage, low-performance cells. Transistor sizing is also exploited to minimize power consumption in combinational logic cells. Rich libraries with many available transistor sizes are very useful in low-power design, because they help synthesis tools in achieving optimum sizing for a wide range of gate loads. Power savings can be obtained by adopting non-standard logic implementation styles such as pass-transistor logic, which can reduce the number of transistors (and, consequently the capacitive load), for implementing logic functions which are commonly found in arithmetic units (e.g., exclusive-or, multiplexers). While most traditional power optimization techniques for logic cells focus on minimizing switching power, circuit design for leakage power reduction is also gaining importance [14].

1.2.2 Analyzed Leakage Current

So far we have discussed power reduction techniques which could be resulted from switching-related currents and current leakage mechanisms. In this section, we will explore the different types of leakage currents and their trends. After that, we will then describe ways to limit various types of leakage.

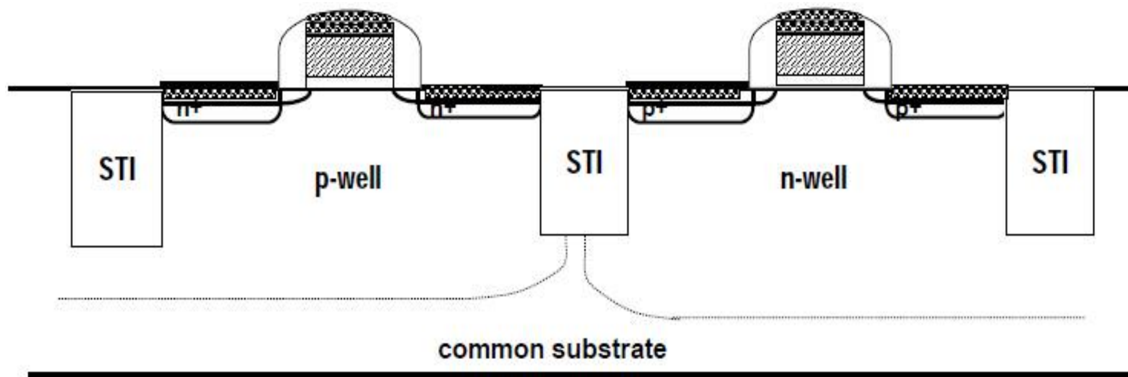


Fig.1.2 Basic CMOS Transistor Structure

Fig.1.2 shows basic CMOS transistor structure. Typical process today uses twin-tub CMOS technology. The main factors to cause change in leakage can be concluded as shallow trench isolation (STI), thin-oxide, lightly-doped drain/source. The main feature of transistors scaling is the reduction in V_{DD}, the threshold voltage (V_t), effective channel length, T_{OX} and doping levels and depth.

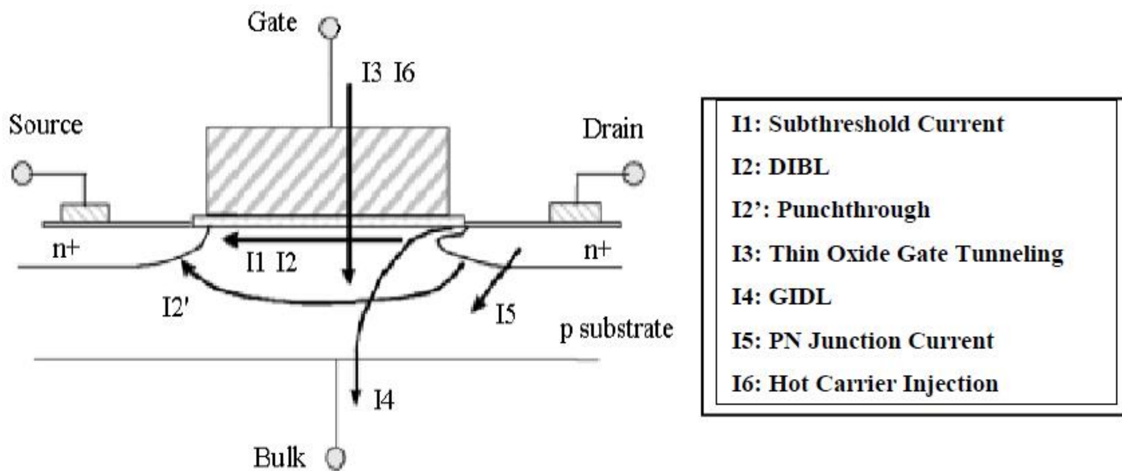


Fig.1.3 Schematic description of the different leakage currents and mechanisms in Twin-tub CMOS transistor.

As analyzed in [15]-[18], the overall leakage currents can be divided into several components (Figure.1.3), taking place under different bias conditions. At very low gate voltage, a potential difference between source

and drain still results in sub-threshold static leakage current, I_1. Among the many parameters, I_1 dependence on higher threshold voltage (V_t) and operation temperature is the most significant, reducing I_1 in an exponential manner with increasing V_t and decreasing temperature, respectively. Basically, lower channel doping, shorter effective channel length and longer transistor width will reduce V_t and increase I_1. In addition, the body-factor and DIBL (Drain-Induced-Barrier-Lowering, I_2) parameters, that depend on the 1D and 2D doping profiles of the V_t adjust halo/pocked and extensions implants will also affect I_1.

Transistor scaling also means shallower and more abrupt extensions and S/D junctions. Although more abrupt junctions provide improved short channel effect, the rising doping concentrations and the high electric field (>10^6V/cm) across the reverse-biased p-n junction lead to leakage due to Band-To-Band-Tunneling (BTBT) [17]. Higher gate-to-drain voltage increases the vertical field in the drain depletion layer, and reduces the depletion width at the gate-drain overlap area, resulting in Gate-Induced-Drain-Leakage (GIDL, I_4) [17]. For having a good V_t control, and to reduce the I_1 leakage, the dopants concentration near the surface are kept high. However, an increase of the drain voltage lowers the potential barrier for the majority carriers at the source side, thus leading in "additional" I_1 leakage and the punchthrough (I'_2).

For technology generation of 65 nm and below, due to aggressive gate oxide thickness (T_{OX}) reduction, direct tunneling through the gate oxide leads to gate leakage (I_{gate}=$I_3 + I_6$), that becomes dominant over I_1. In [16], the gate leakage is simply approximated using W (transistor width) and K_1 and

K_2 that are constants which can be extracted experimentally:

$$I_{gate} = K_1 \times W(\frac{V_{DD}}{T_{OX}})^2 e^{-K_2 T_{OX}/V_{DD}}. \quad \text{(Eq.1.4)}$$

Figure 1.4 describes the gate leakage dependence on the gate oxide thickness. The exponent is much more dominant then the (V_{DD}/T_{OX}) part in the pre-exponent.

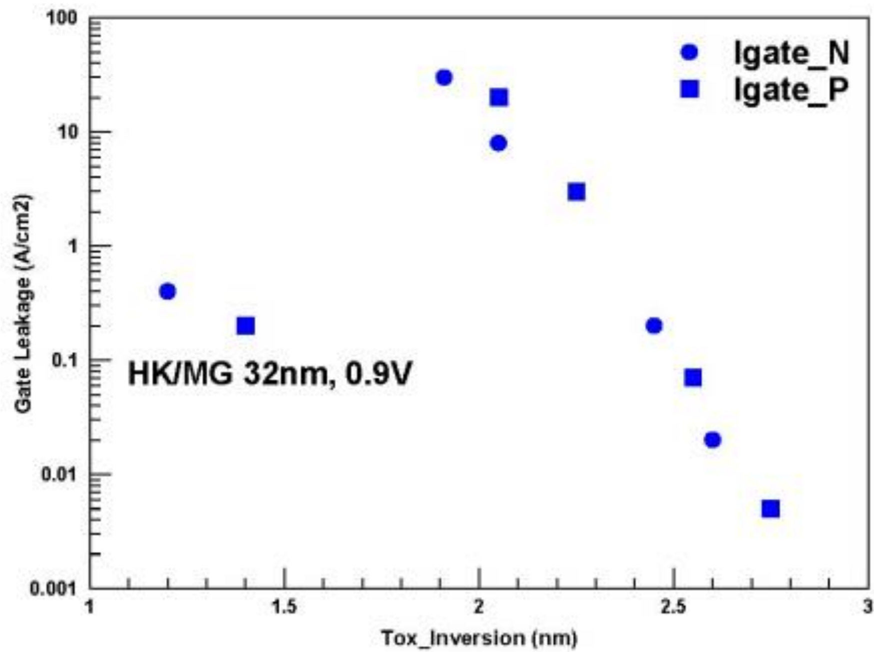


Fig.1.4 Gate leakage vs. Gate oxide thickness for Poly/SiON (65 nm to 40 nm platforms) and HK/MG (for 32 nm). For the same effective oxide thickness, the gate leakage is lower by ~ 3 orders of magnitudes comparing to oxynitridization thermal oxide.

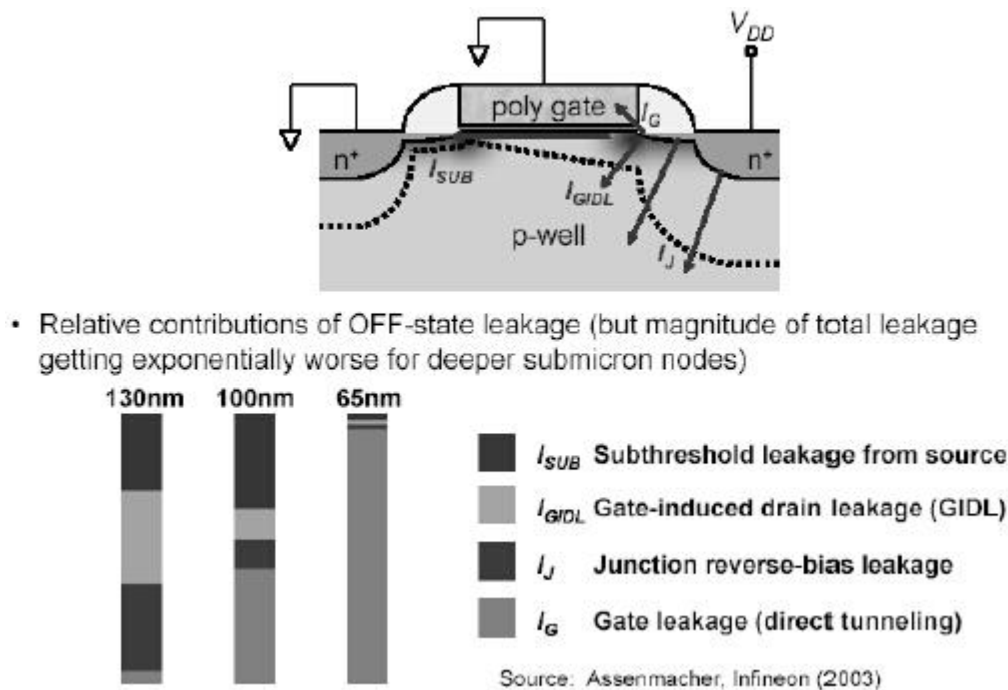


Fig.1.5 Leakage current analysis on different size process

An observation from both Fig.1.4 and Fig.1.5, for 130nm, $\mathbf{I_1}$, GIDL and junction leakage, cover ~95% of the overall leakage, and $\mathbf{I_{gate}}$ < 5%. For 90nm, $\mathbf{I_{gate}}$ is ~40% and for 65 nm, it is >90%. Note that these percentages refer to leakages at room-temperature. As temperature goes-up, both $\mathbf{I_1}$ and the junction leakage become more dominant [19]. Another factor which affects the ratio between the different components is the $\mathbf{V_t}$ target: in multi-$\mathbf{V_t}$ technology, having for example three types of $\mathbf{V_t}$'s, the high-$\mathbf{V_t}$ (H$\mathbf{V_t}$) will have 25% leakage due to $\mathbf{I_{gate}}$, 25% leakage due to diodes and ~50% leakage due to $\mathbf{I_1}$. Regular (or Standard $\mathbf{V_t}$, $\mathbf{SV_t}$) will have <5% for $\mathbf{I_{gate}}$ and diode and ~90% for $\mathbf{I_1}$. In Low-$\mathbf{V_t}$, $\mathbf{I_1}$ is the dominant (>98%) [20]. The 32 nm SL (Standard Logic for General Purposes) foundry technology node is the first one to use high-k material that allows reducing $\mathbf{I_{gate}}$ while keeping good gate control on the channel. About 3

order of magnitude reduction of $\mathbf{I_{gate}}$ can be achieved for the same effective oxide thickness.

In addition to gate current due to tunneling, Hot Carrier Injection (HCI, $\mathbf{I_6}$) at the channel pinch-off area leads to impact ionization and leakage injection into the gate oxide.

Another aspect of scaling is the increase of inter-die thermal gradients due to the increase of the local power densities. Higher thermal gradients increase the voltage drop due to increased leakage. This voltage drop affecting the clock skew. Kawa [21] found voltage drop of 12% and 16% for 30° thermal gradients for 0.18μm and 0.13μm technology nodes, respectively

1.3 Dissertation Organization and Contributions

Based on the previous studies and references of researches, this thesis presents a layout dependent effect-aware leakage current reduction (LDE-aware LCR) technology for ultra-low power and low frequency circuit design.

LDE-aware LCR is used to control the junction leakage, which consumes about 40% energy in 90nm process, more than 90% energy in 65nm process and below.

The junction leakage occurs from the source or drain to the substrate through the reverse-biased diodes when a transistor is OFF. A reverse-biased pn junction leakage has two main components: one is minority carrier diffusion/drift near the edge of the depletion region; the other is due to electron-hole pair generation in the depletion region of the reverse-biased junction [22]. For instance, in the case of an inverter with low input voltage, the NMOS is OFF, the PMOS is ON, and the output voltage is high.

Subsequently, the drain-to-substrate voltage of the OFF NMOS transistor is equal to the supply voltage. This results in a leakage current from the drain to the substrate through the reverse-biased diode. The magnitude of the diode's leakage current depends on the area of the drain diffusion and the leakage current density, which is in turn determined by the doping concentration. If both n and p regions are heavily doped, band-to-band tunneling (BTBT) dominates the pn junction leakage [23]. Junction leakage has a rather high temperature dependency (i.e., as much as 50 – 100°C, but it is generally inconsequential except in circuits designed to operate at high temperatures (> 150°C.) Junction reverse-bias leakage components from both the source-drain diodes and the well diodes are generally negligible with respect to the other three leakage components.

The rest of the dissertation is organized as follows. The detail of LDE-aware LCR is discussed in chapter 2. We analyzed STI induced junction current model. Based on the analysis, we concluded and proposed LDE-aware LCR design procedure. The investigation of single transistor I-V characteristics and simulation of Inverter power consumption support our point of view.

In chapter 3, LDE-aware LCR is used for an ultra-low nano-watt reference circuit design. By using LDE-aware LCR, the transistor to improve performance of sensitive power supply voltage can be saved. Moreover, we propose a novel retargeting methodology by adopting an ACM (advanced compact MOSFET) model to describe the drain current consistently in strong and weak inversion levels. In our case studies, we retarget a circuit from 1.8V/180nm process to 1.2V/90nm, 3.3V/90nm and 1.8V/65nm processes respectively. Besides, we fabricate the reference circuit in the 1.2V/90nm

process, and confirm that good measurement results are obtained with less than 12.8%/V supply voltage variation and only 1.1nW power consumption.

In chapter 4, for the practicality of LDE-aware LCR, we tackle the dedicated leakage reduction technique for the performance of a specific 11-bit SAR ADC. In the practical example of SAR ADC, we discuss the design procedure step by step and implement LDE-aware LCR in fundamental blocks. Our design flow consists of three steps:

- Step 1) is to verify specifications to clarify a proper ADC style.
- Step 2) is to decide an architecture from global view, especially focusing on a switching algorithm and a matching capacitor array.
- Step 3) is to implement LDE-aware LCR and low- power circuit in fundamental blocks.

We verified the power consumption by a triple-well 65nm/1.2V CMOS fabricated SAR ADC. The measured results demonstrate an attractive power consumption of 38.6nW. We also manufacture and measure the same architecture SAR ADC without LDE-aware LCR, and it consumes 55.2nW. That is, introducing our leakage current reduction technology provides a total of 30% energy saving.

Finally, conclusions are given in chapter 5. The previously outlined contributions in chapters 2, 3, and 4, and summarized in the final chapter 5 have resulted in original work published or submitted for publication [24], [25].

References

[1] Alioto, M., "Ultra-Low Power VLSI Circuit Design Demystified and Explained: A Tutorial," Circuits and Systems I: Regular Papers, IEEE Transactions on , vol.59, no.1, pp.3,29, Jan. 2012

[2] "Smartphone Users Worldwide Will Total 1.75 Billion in 2014," WWW.EMARKETER.COM

[3] M. Alioto, Ultra Low-Power/Low-Voltage VLSI Design: From bASICs to Design. Hoboken, NJ: Wiley, to be published.

[4] Benini, L.; De Micheli, G.; Macii, E., "Designing low-power circuits: practical recipes," Circuits and Systems Magazine, IEEE , vol.1, no.1, pp.6,25, First Quarter 2001

[5] "*New Methodologies for Low-Power. High-Performance* Digital VLSI Design," Mohamed W. Allam. A thesis presented to the University of Waterloo in fulfilment of Doctor of Philosophy in Electrical and Computer Engineering

[6] Anantha P. Chandrakasan et al., "Low-Power CMOS Digital Design", IEEE Journal of Solid-State Circuits, vol. SC-27, n. 4, pp. 473-484, 1992.

[7] Technical Support, Pentium III Processor for the SC242 at 450 MHz to 800 MHz, Datasheet, Intel Corporation, Santa Clara, CA, 1999

[8] H.Sasaki, "Multimedia Complex on a Chip", International Solid-State Circuits Conference, pp. 16-19, 1996.

[9] G. K. Yeap, Practical low power digital VLSI design, Kluwer Academic Publishers, Boston, Mass. ; London, 1998.

[10] "Energy Efficiency Trends in Residential and Commercial Buildings", U.S. Department of Energy, Energy Efficiency and Renewable Energy, Oct. 2008

[11] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, “Low-Power CMOS Digital Design”, *IEEE Journal of Solid-State Circuits*, vol. 27, no. 4, pp. 473–484, April 1992.

[12] A. Chandrakasan and R. Brodersen, *Low Power CMOS Design*. IEEE Press, 1998.

[13] J. Montanaro, *et al.*, “A 160-MHz, 32-b, 0.5-W CMOS RISC Microprocessor”, *IEEE Journal of Solid-State Circuits*, vol. 31, no. 11, pp. 1703–1714, November 1996.

[14] Y. Ye, S. Borkar, and V. De, “A New Technique for Standby Leakage Reduction in High-Performance Circuits”, *1998 Symposium on VLSI Circuits*, pp. 40–41, Honolulu, Hawaii, June 1998.

[15] Eitan N. Shauly, "CMOS Leakage and Power Reduction in Transistors and Circuits: Process and Layout Considerations", J. Low Power Electron. Appl. 2012, 2, 1-29.
[16] Helms, D.; Schmidt, E.; Nebel, W. Leakage in CMOS Circuits—An Introduction. In Integrated
Circuit and System Design. Power and Timing Modeling, Optimization and Simulation, 14th
International Workshop (PATMOS 2004); Springer: Berlin, Germany, 2004; pp. 17–35.
[17] Roy, K.; Mukhopadhyay, S.; Mahmoodi-Meimand, H. Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits. Proc. IEEE 2003, 91, 305–327.
[18] Mann, R.W.; Abadeer, W.W.; Breitwisch, M.J.; Bula, O.; Brown, J.S.; Colwill, B.C.; Cottrell,
P.E.; Crocco, W.G., Jr.; Furkay, S.S.; Hauser, M.J.; et al. Ultralow-power SRAM technology. IBM J. Res. Dev. 2003, 471, 553–563.
[19] Tavel, B.; Duriez, B.; Gwoziecki, R.; Basso, M.T.; Julien, C.; Ortolland, C.; Laplanche, Y.; Fox, R.; Sabouret, E.; Detcheverry, C.; et al. 65 nm LP/GP Mix Low Cost Platform for Multi-Media Wireless and Consumer Applications. In Proceedings of the 35th European Solid-State Device Research Conference (ESSDERC 2005), Grenoble, France, 12–16 September 2005; Volume 50, pp. 573–578.
[20] Skotnicki, T.; Fenouillet-Beranger, C.; Gallon, C.; Boeuf, F.; Monfray, S.; Payet, F.; Pouydebasque, A.; Szczap, M.; Farcy, A.; Arnaud, F.; et al. Innovative materials, devices, and CMOS technologies for low-power mobile multimedia. IEEE Trans. Electron Device 2008, 55, 96–130.
[21] Kawa, J. Low power and power management for CMOS—An EDA perspective. IEEE Trans.
Electron Device 2008, 55, 186–196.
[22] B. Van Zeghbroeck, Principles of Semiconductor Devices, http://ece. www.colorado.edu/~bart/book/book/title.htm, ch.4.
[23] Y. Taur and T. H. Ning, Fundamentals of Modern VLSI Devices. New York: Cambridge Univ. Press, 1998, ch. 2, pp. 94–95.
[24] Gong Chen; Bo Yang; Nakatake, S.; Zhangcai Huang; Inoue, Y., "A retargeting methodology of nano-watt CMOS reference circuit based on advanced compact

MOSFET model," Circuits and Systems (ISCAS), 2012 IEEE International Symposium on , vol., no., pp.938,941, 20-23 May 2012

[25] Gong Chen; Zhang Yu; Bo Yang; Qing Dong; Nakatake, S., "A comparator energy model considering shallow trench isolation stress by geometric programming," Quality Electronic Design (ISQED), 2013 14th International Symposium on , vol., no., pp.585,590, 4-6 March 2013

Chapter -2 LDE-aware Leakage Current Reduction Technique

2.1 Preliminary Investigation

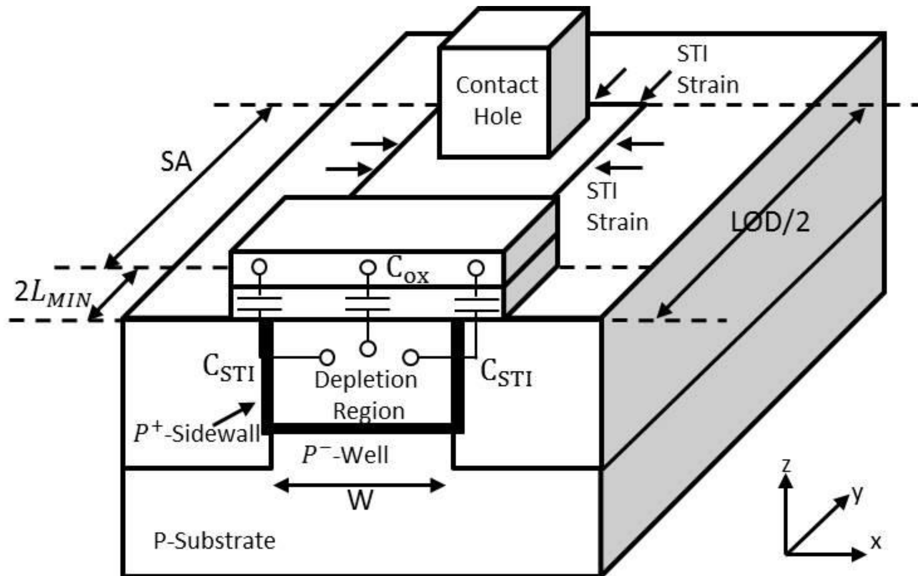
There are several works addressing the STI issues. In [26], author proposes a complete flow to characterize the influence of STI stress on performance in RF/analog circuits based on the layout design and process information. A model for STI is presented in [27], [28] and [29] contributing to the mobility, velocity saturation, threshold voltage and changes of some second-order effects, such as drain induced barrier lowering (DIBL) and the body effect. A process-aware channel edge shaping method is proposed in [30], [31] and [32] to reduce the leakage current in the STI.

Generally, the issue of leakage current reduction technology is often discussed in DRAM or SRAM systems. A suppression methodology of edge effect of leakage current reduction in sub-40nm DRAM device is proposed in [32]. A high density low-power full CMOS SRAM cell technology with an STI is proposed in [33]. The STI induced junction leakage mechanism effects on memory devices are also discussed in [34] and [35]. Most of the leakage current reduction technologies are related to the LDE or process parameters.

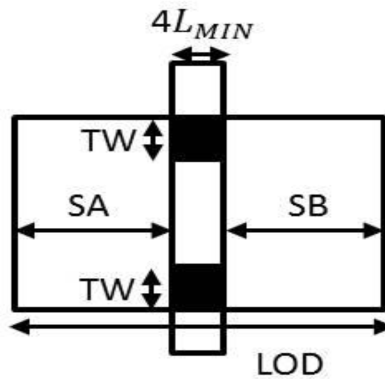
The STI induced leakage current introduces both power inefficiency and logic function errors on memory devices. Because the switch capacitors have enough time to discharge, for low-speed SAR ADC, leakage current caused charge injection would cause less function errors or performance degradation. However, the leakage current contributes to be an influencing factor in power consumption. So far, researches for leakage current reduction

in ADCs still have not been explored in depth. The measured results suggest that the LDE-aware LCR for low-speed high power efficient SAR ADC is as important as some other essential procedures, such as the common-centroid layout design.

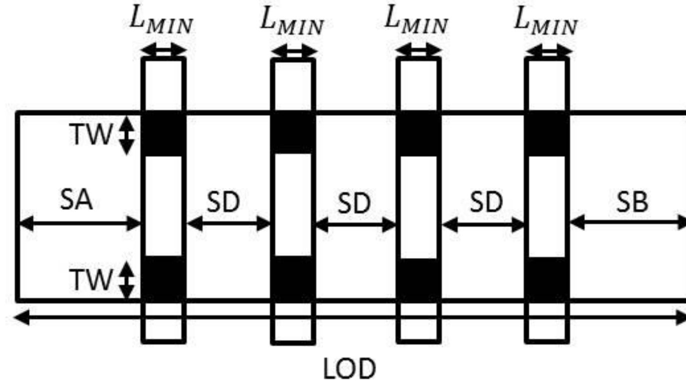
2.2 STI Specifications



(a) Cross section of a single finger STI-isolated MOSFET



(b) Top view of single finger MOSFET



(c) Top view of multiple finger MOSFET

Fig. 2.1 STI-isolated MOSFET

Parameter	Value
S/D depth	0.1-um
S/D length	0.5-um
S/D doping	$2.1 \times 10^{20} cm^{-3}$
Parasitic MOSFET width	1-um
P^{+} sidewall doping	$8.2 \times 10^{17} cm^{-3}$
P^{-} well doping	$5.5 \times 10^{16} cm^{-3}$

Table2.1 TECHNOLOGY AND GEOMETRIC PARAMETERS OF STI STRUCTURE

Summarizing the preliminary methods, we propose an LDE-aware LCR to implement low-power in the SAR ADC analog blocks. STI stress, which is exerted by STI wells on an active region of the device, is inevitably formed and has increasingly significant impact on device behavior. Fig. 1 demonstrates the cross section and top views of the MOSFET STI model. The model has two structures, one is a single finger MOSFET, and another is multiple finger MOSFET. The device structure in Fig. 1 (a) identified as a

two dimensional of lateral (x-) and vertical (y-) parasitic MOSFET, which is formed parasitic capacitances between gate, STI sidewall, and active area. The significant leakage current across STI region when sidewall guard p_{-} doping cannot properly suppress the inversion channel formation due to the impact of STI interface-trapped charges. Basically, a commercial 3D TCAD process and device simulator, such as Nova TCAD, can simulate the junction leakage by tuning the technology and geometric parameters in Table 1. However, for engineers, only geometric size, such as S/D length, parasitic MOSFET width are the possible tuning parameters.

Rule	Design Rule Description	180 nm	130 nm	90 nm	65 nm
GC.D.1	Distance of Poly (over STI) to related AA	0.10	0.07	0.05	0.05
GC.X.2	Extension of poly beyond AA (end-cap)	0.22	0.18	0.16	0.14
AA.D.3	Distance between WN to N+ in WP	0.43	0.34	0.22	0.16
AA.E.3	Enclosure of WN around P+ in WN	0.43	0.34	0.22	0.16
CS.D.1/2	Distance of CS over AA to related Gate	0.16	0.11	0.11	0.09

Table2.2 Main design-rules needed for building transistors in standard-cell library. See Fig.2.2 for illustration of these rules.

An example for layout of single finger MOSFET including STI is shown in Fig2.2. For drawing a typical 3 input NOR gate, about 20 TDR (Topological Design Rules) are needed. Among them, several rules have a direct relation with the transistor leakage, and therefore, should be optimized for low-power design. The analysis below covers some of these layout rules that are listed in Table2.2. From Fig2.2, considerations for stress affecting layout modification are:

- Expending the active area (AA, Source and Drain) edges beyond gate (GC.X.1), for different STI induce stress (Figure 2.3). The stress range and magnitude are up to GC.X.1 = 1.3μm, and <10%, respectively;
- Re-placement of contacts with distance to gate (CS.D.1). This is because contacts "punching" of the cSEL layer, and release some of the stress. For this reason, also re-set of the Source and Drain contact pitch may improve performance [36];

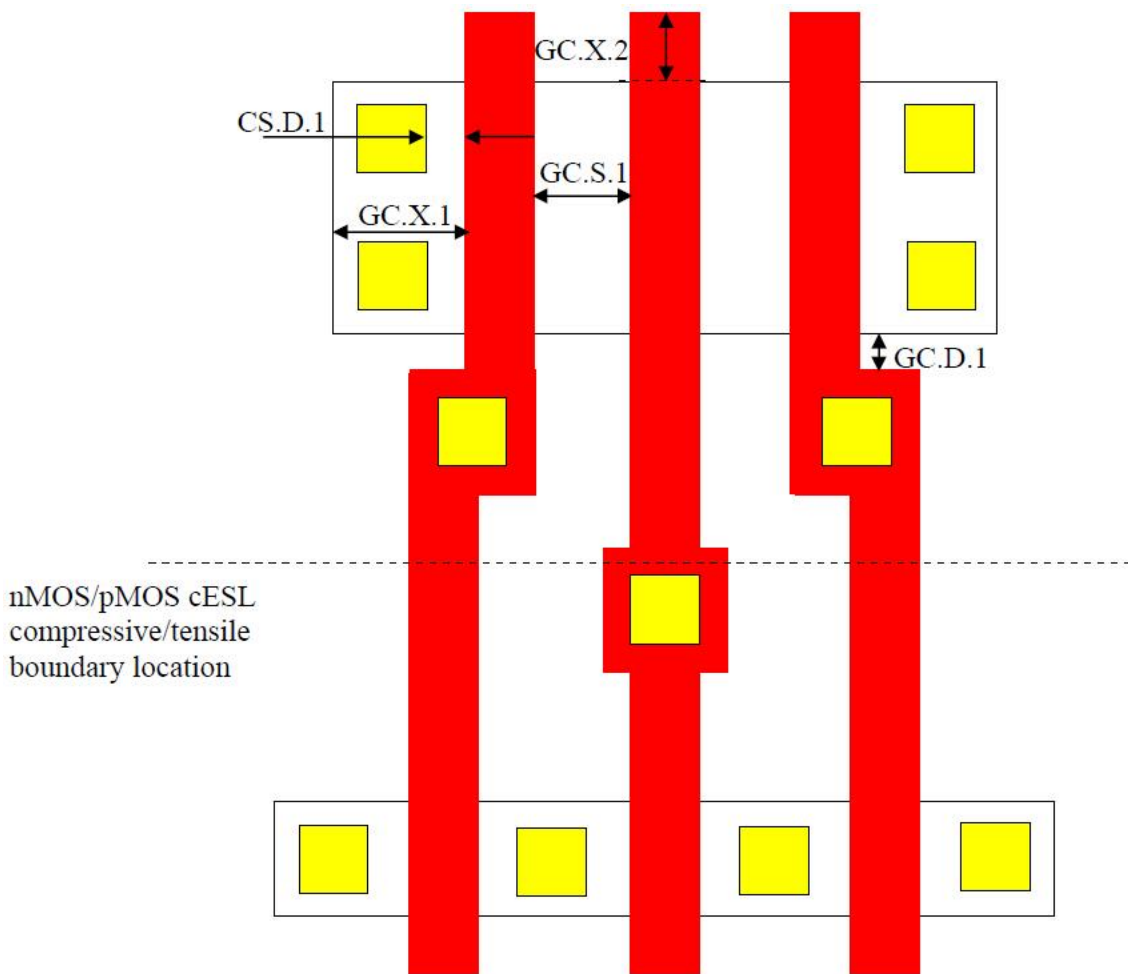


Fig.2.2 Typical 3 input NOR gate with some of different Topological Design Rules.

- Poly space between transistor fingers (without contacts) [37]. This is because smaller space also means narrow cSEL layer, or narrow eSiGe stressor trench, or both. In [38], up to 7.8% degradation was seen for different poly spaces;
- Location of tensile/compressive nitride cSEL boundary layer over STI (Fig.2.2) and separating nMOSFET and pMOSFET [37].

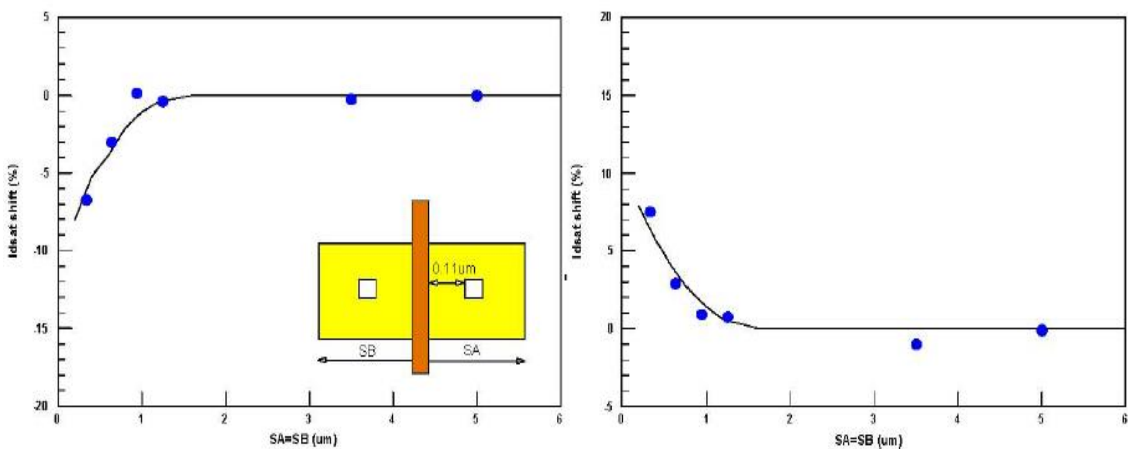


Fig.2.3 The effect of AA extension beyond gate (GC.X.1) LOD stress and on drive current for (left) nMOSFET, (right) pMOSFET. For both devices, W = 1.2 μm, L = 0.13 μm, Single finger, and SA = SB, where SA and SB are the extensions from left and right of the transistor, respectively.

2.3 STI Induced Junction Current Model

The S/D length modify is introduced in BSIM4 [27], [28]. An STI stress model which follows the 1/LOD trend but reveals differences L and W

scaling had been developed. In this BSIM4 model, the LOD is basically modulated through varying finger number (NF), SA, SB and SD, as shown in Fig. 1 (b) and (c). Hereinafter, SA is the distance between OD edge to ploy from one side, SB is the distance between OD edges to ploy from the other side. SD is defined as the distance between OD neighboring fingers. SA is set to equal SB for the symmetrical consideration. SD is a constant which equals the minimum distance of design rule requirement.

According to the MOSFET off-current equation is:

$$I_{leakage} = I_S(V_{GS} = 0V) = I_0 \frac{W}{L} \exp(-\frac{V_{th}}{\eta V_T}) \quad \text{(Eq.2.1)}$$

The minimum leakage current $I_{leakage}$ is either caused by an V_{th} increase or an improved sub-threshold slope η. The MOSFET threshold voltage equation is

$$V_{th} = V_{th0-original} + \gamma\left(\sqrt{2\phi_0 + V_{SB}} - \sqrt{2\phi_0}\right) \quad \text{(Eq.2.2)}$$

$$V_{th0-original} = V_{FB} + \phi_0 + \triangle V_{SCB} + \frac{Q_B}{C_{OX}} \quad \text{(Eq.2.3)}$$

Here, ϕ_0 is the surface potential, V_{FB} is the flat-band voltage, Q_B is the body charge, C_{OX} is the oxide capacitance and γ is the body-effect parameter. The STI introduce novel variations of S/D diffusion length into the Eq.2.2 and Eq.2.3. The STI-aware threshold voltage expression is:

$$V_{th0} = V_{th0-original} + \frac{KV_{th0}}{K_{stress_{vth0}}}(Inv_{SA} + Inv_{SB} - Inv_{SAref} - Inv_{SBref}) \quad \text{(Eq.2.4)}$$

Where KV_{th0} is the threshold shift parameter for stress effect, $K_{stress_{vth0}}$ is a process P-type well (PW) misalignment produced variable, which can be roughly regarded as a constant value. SAref is the reference distance between OD and edge to poly of one side. SBref is the reference distance between OD and edge to poly of the other side. Theoretically, to choose the SAref and SA

has the same length can eliminate the STI induced threshold variable. Hereinafter, all simulations are set to have the same SAref and SA length. For single finger MOSFET, $\boldsymbol{Inv_{SA}/Inv_{SB}}$ is:

$$\boldsymbol{Inv_{SA} = Inv_{SB} = \frac{1}{SA+0.5L_{drawn}}} \quad \text{(Eq.2.5)}$$

And for multiple finger MOSFET is:

$$\boldsymbol{Inv_{SA} = Inv_{SB} = \frac{1}{NF}\sum_{i=0}^{NF-1}\frac{1}{SA+0.5L_{drawn}+i(SD+L_{drawn})}} \quad \text{(Eq.2.6)}$$

Where $\boldsymbol{L_{drawn}}$ is the actual gate length.

Except the S/D length tuning, parasitic MOSFET dimension adjusting also contributes to the leakage current performance. As reported in [30], [31] and [32], the abrupt transient portion from the isolation to the active region causes the impurity segregation and fringe electrical field near the active edge. These edge effects lead to the local decrease in the threshold voltage resulting in the increase of the leakage current near the active edge. To reduce the edge effects, partial gate channel dimension is modified. As seen in Fig.2. 1 (b) and (c), the gate channel is divided into the parasitic region operation in parallel of the ploy gate surface (black portion). And the intrinsic region forms at the center (white portion). The leakage current degrades with setting the longer parasitic region than the channel length L. By defining the width of parasitic region is TW, the leakage current per unit width can be estimated as:

$$\boldsymbol{I_{leakage}(W,L) = F(W,L,x)}$$

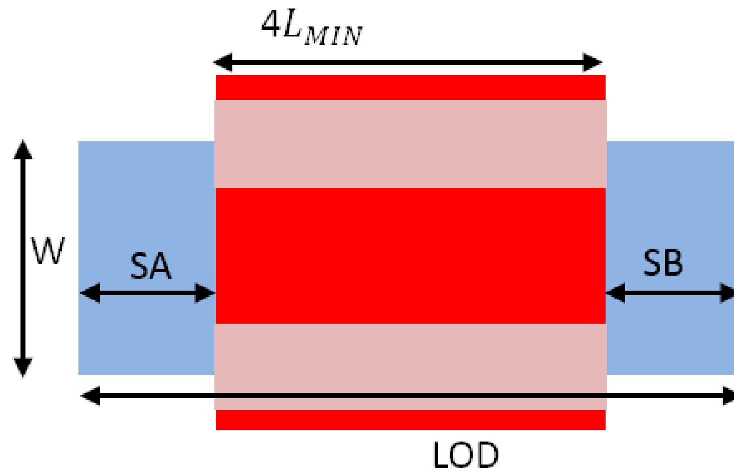
$$\boldsymbol{= I_{leakage-c}(L) \times W + \int_{x=0}^{TW} I_{leakage-f}(x,L)}$$

$$\boldsymbol{=I_{leakage-c}(L) \times W + \int_{x=0}^{TW} I_{leakage-f}(0,L) \times I_{leakage-f}(x)} \quad \text{(Eq.2.7)}$$

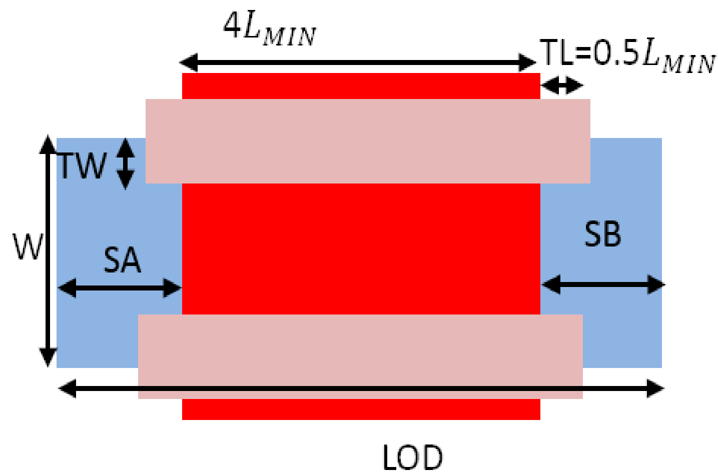
$\boldsymbol{I_{leakage-c}}$ and $\boldsymbol{I_{leakage-f}}$ represents the leakage current per intrinsic region and unit width at the parasitic region, respectively. $\boldsymbol{I_{leakage-f}(0,L)}$

is the leakage current at the active edge (x=0) of given L. $\boldsymbol{I_{leakage-f}(x)}$ means the variation of $\boldsymbol{I_{leakage-f}(0, L)}$ according to the increase of x.

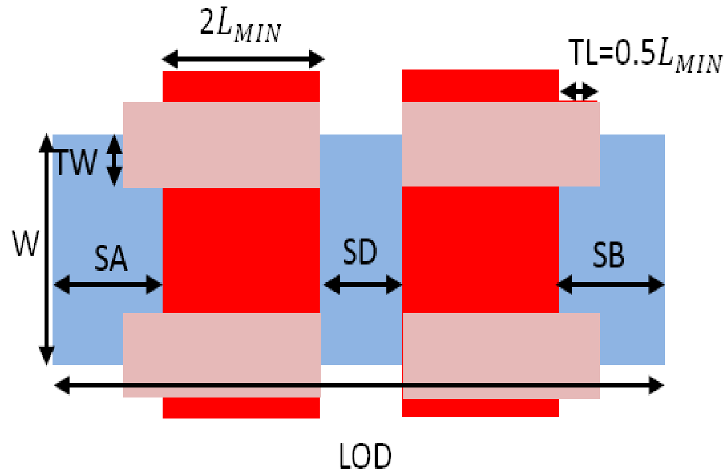
2.4 Descriptions of Technique of LDE-aware LCR



(Step 1) MOSFET layout by considering SA=SB Design



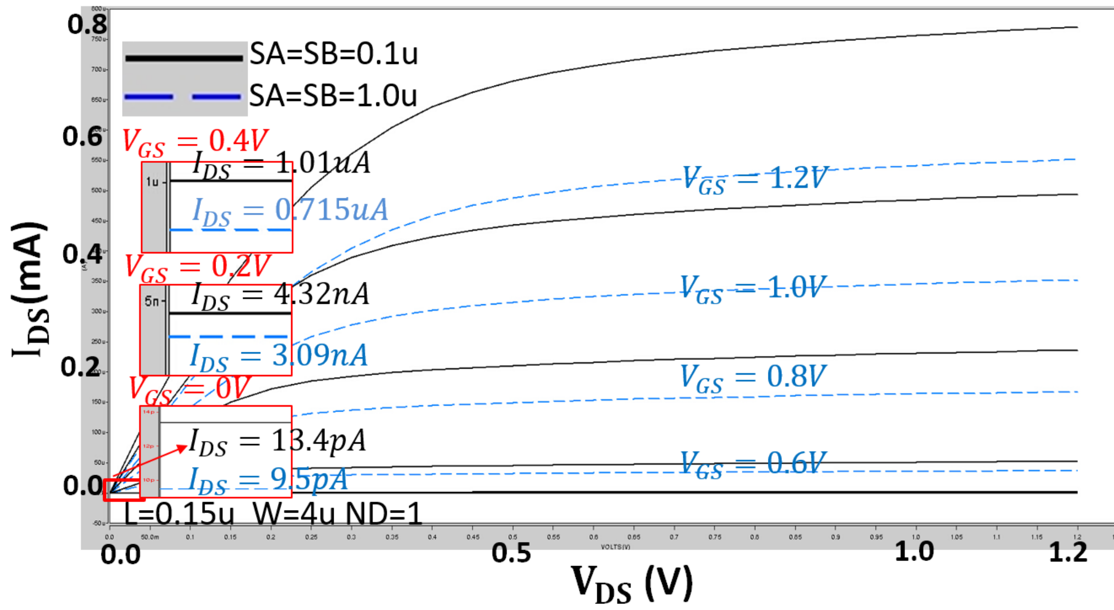
(Step 2) MOSFET layout by considering gate channel shaping (GCS) Design



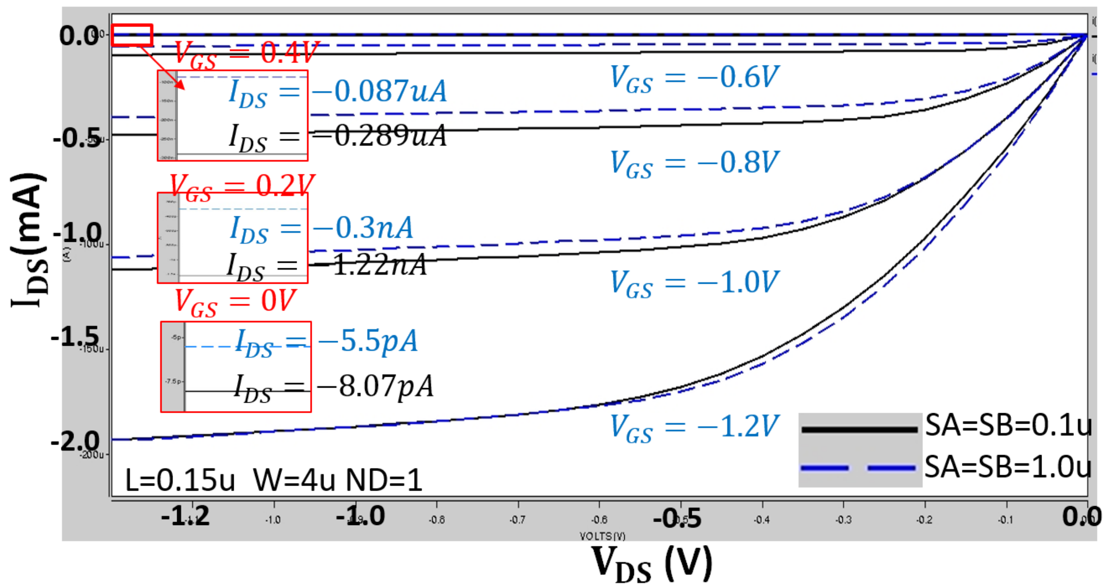
(Step 3) MOSFET layout by considering multiple fingers

Fig. 2.4 Proposed LDE-aware LCR Design Procedure

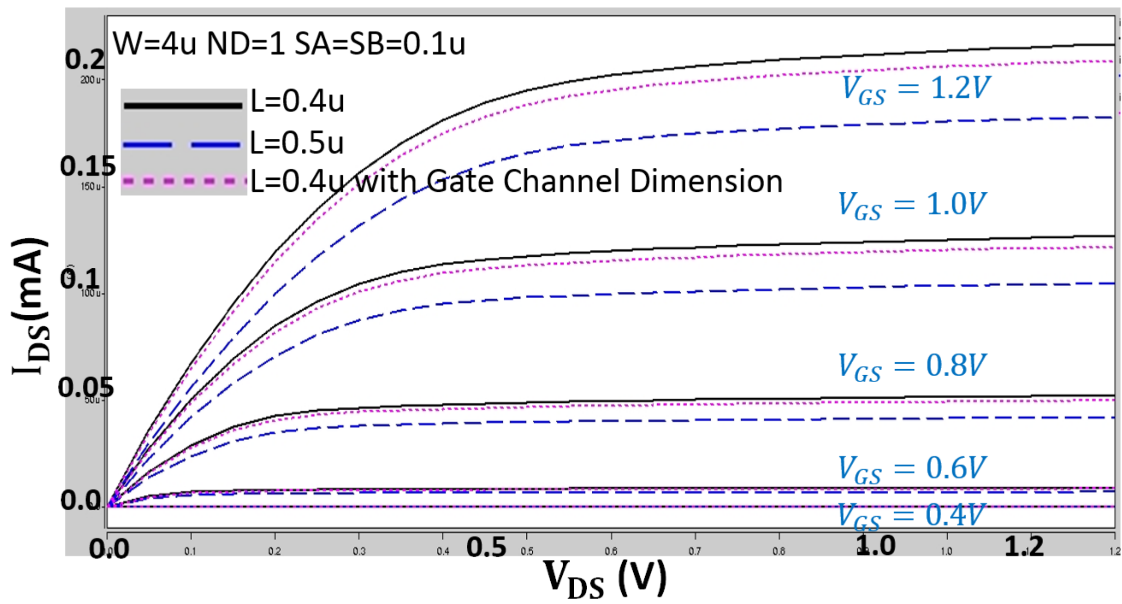
Based on the previous analysis, we conclude that the circuit designer should consider three steps for the limitation of leakage current. As shown in Fig.2.4, the first step is to determine the SA=SB length. In the single transistor I-V characteristics simulation (65nm process) results of Fig.2.5 (a) and (b), it is obvious that both PMOS and NMOS have larger $\boldsymbol{I_{DS}}$ when SA=SB=0.1u. In the detail of $\boldsymbol{V_{GS}}$= 0. 4V, in which the transistor works in the weak inversion region, the current shows about 0.2 μA difference. In Fig. 2.4, the second step is to implement gate channel shaping (GCS), and the third step is to determine the finger numbers. Also Fig. 2.5 (c), (d), (e) and (f) verify the validation of these two steps. For the DRC check consideration, it is noted that the implementation of GCS should under condition of transistor width exceeding twice of minimal width. It means that the second step of LDE-aware LCR should with carefully applied in the digital circuit layout design, which are always following the minimal width rules in the standard cell library.



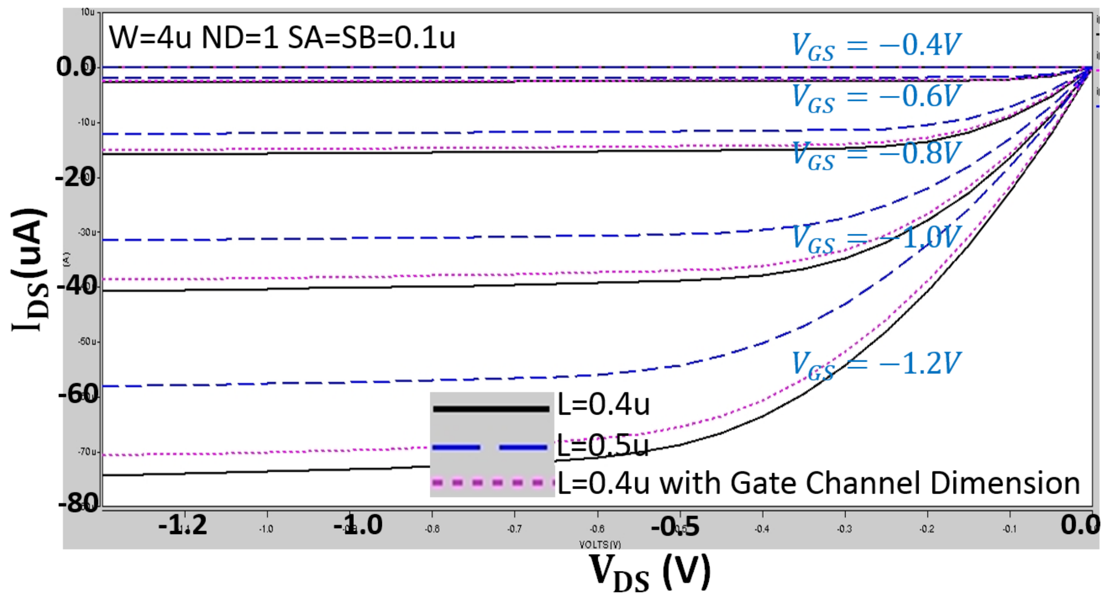
(a) SA=SB effects on Single NMOS



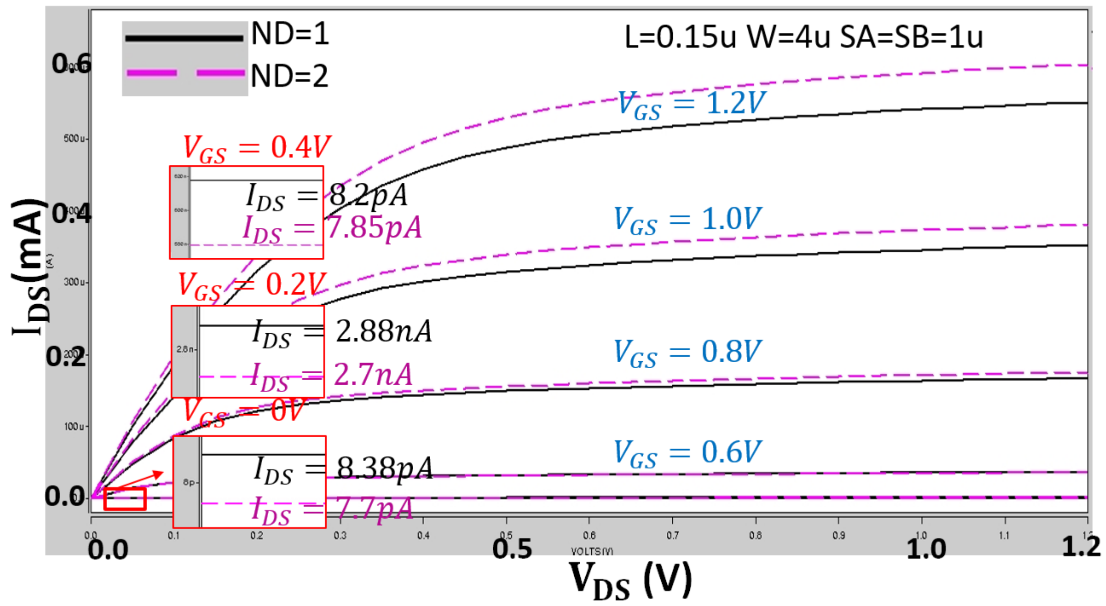
(b) SA=SB effects on Single PMOS



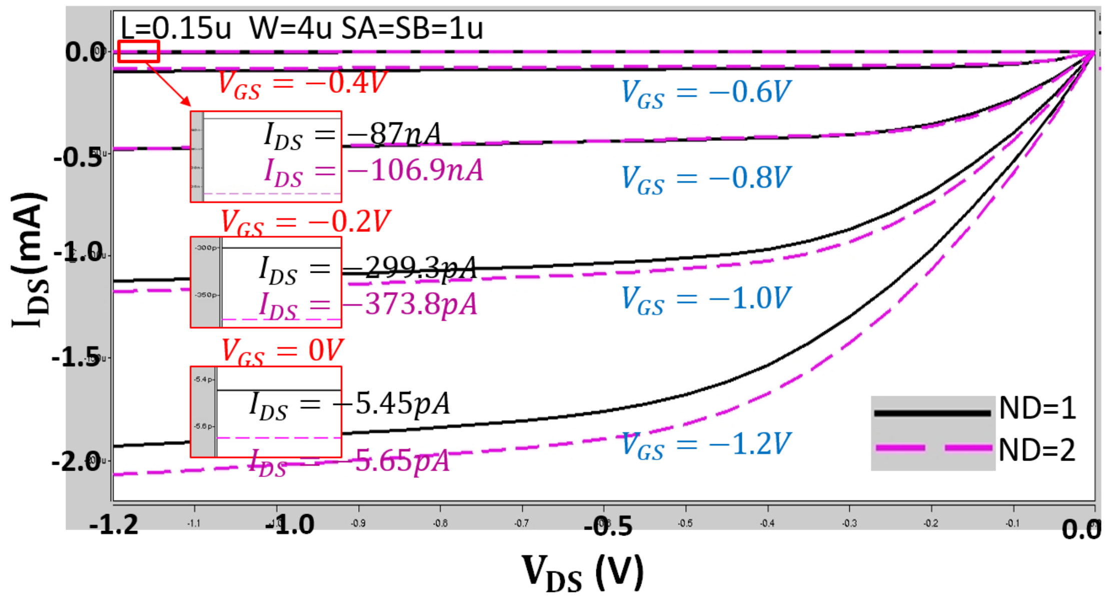
(c) GCS effects on Single NMOS



(d) GCS effects on Single PMOS



(e) Multiple fingers effect on Single NMOS



(f) Multiple fingers effect on Single PMOS

Fig.2.5 The validation of LDE-aware LCR based on the I-V characteristics investigation in 65-nm CMOS

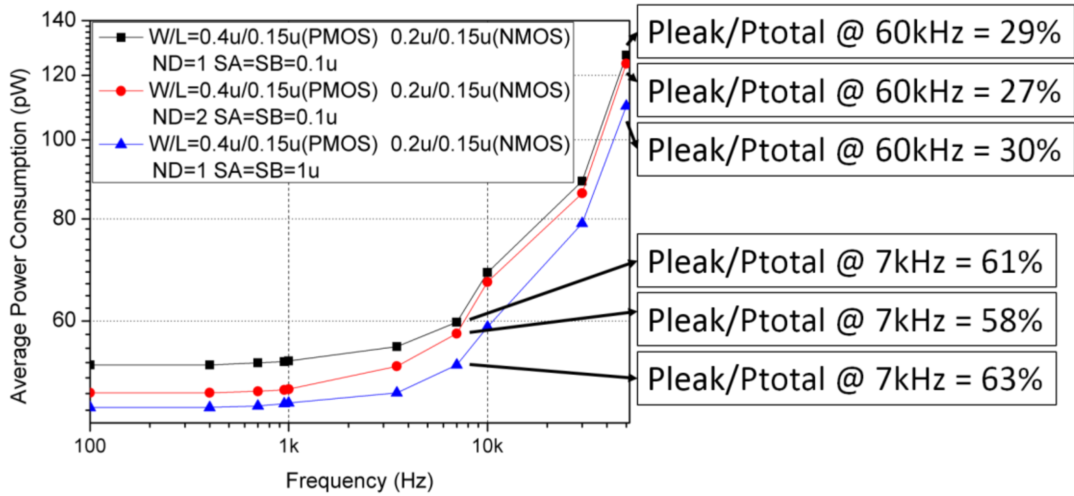


Fig.2.6 Simulated average power consumption versus switching frequency of an inverter with a fan-out of four

Fig.2.6 demonstrates the average power consumption of an inverter as a function of its switching frequency. The power consumption is simulated with three different setting but with same transistor size, the dimension of NMOS is $2.5Wmin/2.3Lmin$ = 0.2μ/0.15μ, and PMOS is $5Wmin/2.3Lmin$ = 0.4μ/0.15μ. The length of transistors are chosen to meet ND=2 simulation. Noticed that the leakage power at 1-7 kHz can constitute more than 58% (58% at 7 kHz) of the total power. In the higher frequency of 60 kHz, the leakage power accounts for about 27%of all energy consumption. The proof of inverter supports our conclusion that the leakage currents contributes to a significant portion of the total power consumption, and proves that the total power dissipation is different with particular LDE parameters. Thus our method of LDE-aware LCR for total power optimization becomes consequential.

References

[26] Jiying Xue; Zuochang Ye; Yangdong Deng; Hongrui Wang; Liu Yang; Zhiping Yu, ”Layout-dependent STI stress analysis and stressaware RF/analog circuit design optimization,” Computer-Aided Design - Digest of Technical Papers, 2009. ICCAD 2009. IEEE/ACM International Conference on , vol., no., pp.521-528, 2-5 Nov. 2009

[27] BSIM4.5 Compact Model, www.berkeley.edu.

[28] BSIM4.6 Compact Model, www.berkeley.edu.

[29] CMC Website: www.eigroup.org/cmc.

[30] Oishi, T.; Shiozawa, K.; Furukawa, A.; Abe, Yuji; Tokuda, Y., ”Isolation edge effect depending on gate length of MOSFETs with various isolation structures,” Electron Devices, IEEE Transactions on , vol.47, no.4, pp.822,827, Apr 2000

[31] P. Gupta et al., ”Shaping gate channels for improved devices”, Proc. SPIE, vol. 6925, p. 6925OI, 2008

[32] Soo-Han Choi; Young Hee Park; Chul-Hong Park; Sang-Hoon Lee; Moon-Hyun Yoo; Kim, Gyu-Tae, ”Efficient characterization and suppression methodology of edge effects for leakage current reduction of sub-40nm DRAM device,” Microelectronic Test Structures (ICMTS), 2010 IEEE International Conference on , vol., no., pp.34,37, 22-25 March 2010

[33] Soon Moon June; Sung-Bong Kim; Jung-Sup Uom;Won Suek Cho; Joo Young Kim; Kyung Tae Kim, ”High density low power full CMOS SRAM cell technology with STI and CVD Ti/TiN barrier metal,” VLSI and CAD, 1999. ICVC ’99. 6th International Conference on , vol., no., pp.119,121, 1999

[34] Ha, D.; Changhyun Cho; Shin, D.; Gwan-Hyeob Koh; Tae-Young Chung; Kim, Kinam, ”Anomalous junction leakage current induced by STI dislocations and its impact on dynamic random access memory devices,” Electron Devices, IEEE Transactions on , vol.46, no.5, pp.940,946, May 1999

[35] Maji, D.; Liao, P.J.; Lee, Y.-H.; Shih, J.R.; Chen, S.C.; Gao, S.H.; Lee, J.H.; Wu, K., ”A junction leakage mechanism and its effects on advance SRAM failure,” Reliability Physics Symposium (IRPS), 2013 IEEE International , vol., no., pp.3E.1.1,3E.1.5, 14-18 April 2013

[36] Ban, Y.; Pan, D.Z. Compact Modeling and Robust Layout Optimization for Contacts in Deep
Sub-wavelength Lithography. In Proceedings of the Design Automation Conference (DAC), Anaheim, CA, USA, 13–18 July 2010.
[37] Chakraborty, A.; Shi, S.X.; Pan, D.Z. Layout Level Timing Optimization by Leveraging Active
Area Dependent Mobility of Strained-Silicon Devices. In Proceedings of the Design, Automation & Test Europe (DATE), Munich, Germany, 10–14 March 2008.
[38] Miyashita, T.; Ikeda, K.; Kim, Y.S.; Yamamoto, T.; Sambonsugi, Y.; Ochimizu, H.; Sakoda, T.; Okuno, M.; Minakata, H.; Ohta, H.; et al. High-Performance and Low-Power Bulk Logic Platform Utilizing FET Specific Multiple-Stressors with Highly Enhanced Strain and Full-Porous Low-k Interconnects for 45-nm CMOS Technology. In Proceedings of the IEEE International Electron Devices Meeting, (IEDM 2007), Washington, DC, USA, 10–12 December 2007; pp. 251–254.

Chapter -3 Nano-watt CMOS Reference Circuit Retargeting Methodology with LDE-aware LCR

Focusing on a nano-watt CMOS reference circuit, we propose a novel retargeting methodology by adopting an ACM (advanced compact MOSFET) model to describe the drain current consistently in strong and weak inversion levels. Taking the effect of very long channels into account, we formulate the threshold voltage as a function of the drain-source voltage. While supply voltage is changing, the variable of the drain-source voltage causes reference current and reference voltage outputs at various tilts. Thus redesigning geometric size of transistors is required for different process. Based on the ACM model, we approximate all base conditions in terms of ratios of the channel widths and lengths. From the derivation, we define a tuning parameter with an empirical range and fix all transistor sizes by sweeping this parameter value as well as applying a computer aided design. In our case studies, we retarget a circuit from 1.8V/180nm process to 1.2V/90nm, 3.3V/90nm and 1.8V/65nm processes respectively. Besides, we fabricate the reference circuit in the 1.2V/90nm process, and confirm that good measurement results are obtained with less than 12.8%/V supply voltage variation and only 1.1nW power consumption.

3.1 Preliminaries

A well-known circuit with MOSFET transistors and only one resistor was proposed in [39]. Many voltage references use it to generate the PTAT current because the circuit is very simple and can work in very low power supply voltage. However, this circuit has a drawback for some applications

because of the resistor. Therefore, the authors in [40] proposed a current reference circuit without resistors based on the same structure as [39]. However, it has the weak points with low PSRR (10% per V) [40].

Then, an improved circuit was proposed by the authors in [41]. The sensitivity of its reference current to the supply voltage is relatively low compared with [40]. Moreover, the circuit in [10] is very attractive because it has extremely low power consumption with only several nano-watt. However, the temperature coefficient of the circuits in [40] [41] are still not very low.

Finally, in [42], a reference circuit, which is also efficient and easy-to-design similar as [40] [41], is proposed by adding three MOS transistors based on the circuit [10]. As shown in Fig.3.1, the leakage compensation technique is utilized to improve the temperature performance. Except the extra functional transistors of M3, M5 and M10, the reference circuit can be divided into two functional modules, namely, self-cascode MOSFET (SCM) and selfbiased circuit (SBC) as shown in Fig. 2. These two portions are widely used in various kinds of reference circuits. The analysis of SCM and SBC has universal significance for other reference circuits.

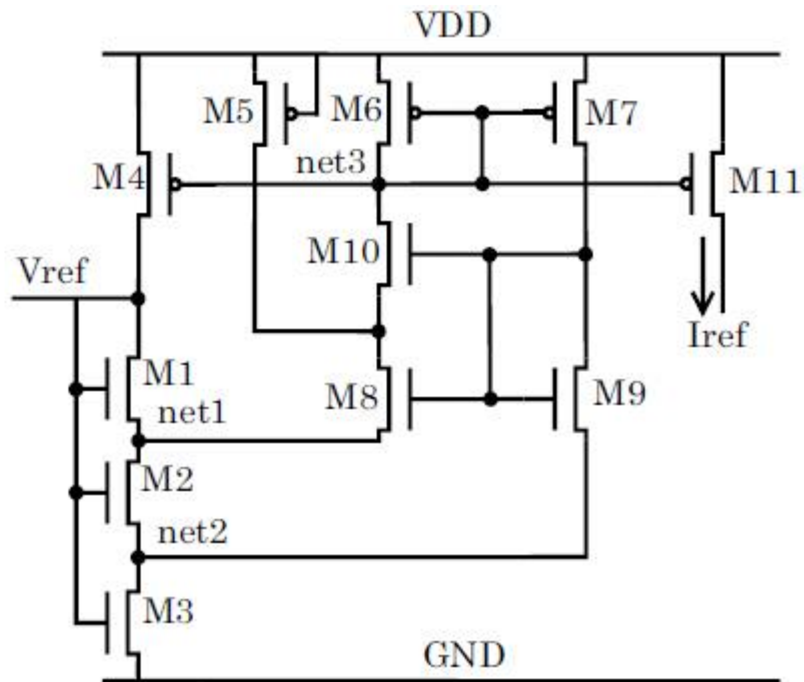


Fig.3.1 CMOS reference circuits in [42]

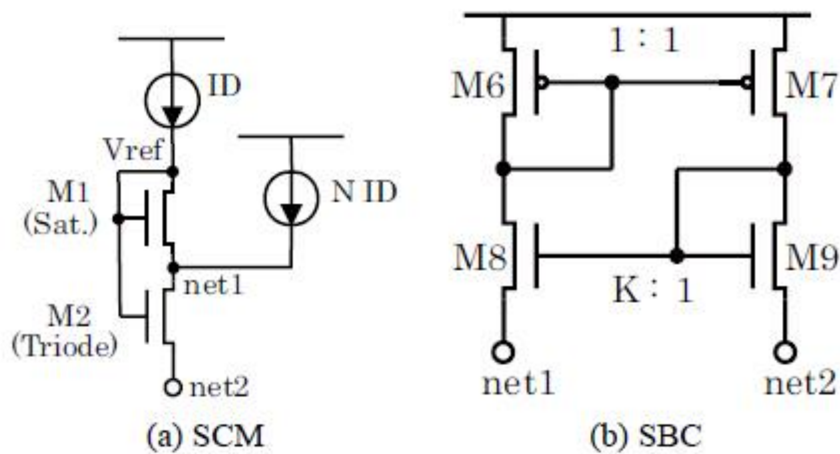


Fig.3.2 Schematics of the essential circuits

In the other hand, in general, a large analog system is composed of many fundamental circuits, the reusability of fundamental circuits is a key issue to improve the design productivity of analog systems. In addition, low

power consumption reference circuits are widely used in circuit systems, their reusability has become an issue in nowadays analog design process.

One of the frequent cases to reuse circuits is retargeting to a different manufacturing process. There have been studies [43]–[46] of several CMOS circuit models taking the retargeting into account. However, conventional analog scaling techniques [47], [48] to conserve bias currents only in the strong inversion level cannot be applied any longer because low power circuits aggressively use the weak inversion of MOSFETs. Furthermore, huge channel lengths of MOSFETs are used for providing slight leakage currents, and it complicates the retargeting due to the dependency of the threshold voltage on the drain-source voltage.

3.2 A conventional method to derive reference current and voltage based on re-defined V_{th}

In the beginning, some conventional methods are applied to the process migration of the reference circuit. Scaling factors are provided in previous work [47], [48] for retargeting from the original process to the target process. Unfortunately, in our case study to apply the above scaling factors to the SBC circuit, the simulation did not output a good result. Both the circuit performance and the value of reference current and reference voltage do not meet the required specifications.

The reason that causes the problem is the use of long channel length transistors. The dependency of V_{th} on the channel length influences the sensitivity of V_{ref} and I_{ref} when V_{DD} changes. In order to see the influence visually, it is necessary to re-define V_{th} based on the level 54 BSIM MOS model in [49].

The reference current described in [42] is as follows:

$$I_{ref} = \frac{1}{2}k' S_{M1} P_2^2 [P_1 + \sqrt{1 + P_1^2}]^2. \quad \textbf{(Eq.3.1)}$$

Where $P_1 = \frac{S_{M2}}{S_{M1}} \frac{S_{M5}}{S_{M5}+S_{M6}}$. $P_2 = nV_T ln(\frac{S_{M6}S_{M8}}{S_{M7}S_{M9}})$. And n is the subthreshold slop factor. $V_T = kT/q$ is the thermal voltage. All $S_M s$ are the *W*/*L* ratios of the transistors referred in Fig.3.1, *W* and *L* are the transistor length and width. k' is the transconductance parameter of MOS transistors. This equation is calculated under the condition that M1, M2, M3, M8 and M9 have the same V_{th}, thus I_{ref} is only associated with the geometrical size of transistors. As a result, migration only changes the value of I_{ref}.

We re-calculate V_{ref} and I_{ref} without neglecting V_{th}. In the level 54 BSIM MOS model [49], the V_{th} dependency on the technological parameters and bias is given as:

$$V_{th} = V_{fb} + V_{phi} + B_1 V_{bs1} - B_2 V_{bs2} - B_{eta} V_{ds}. \quad \textbf{(Eq.3.2)}$$

Where V_{fb} is the flat band voltage and V_{phi} is the surface potential. $B_1 V_{bs1}$ and $B_2 V_{bs2}$ are the root subthreshold voltage and the linear subthreshold voltage, respectively. All MOSFET model parameters are defined as monomials. V_{ds} is related to V_{th} as seen in $B_{eta} V_{ds}$. B_{eta} is a process parameter that includes three physical impacts: the static feedback factor for adjusting the V_{th} due to the difficulty of band bending; the drain-induced barrier lowering (DIBL) effect coefficient for the V_{th}; and the channel-length independent drain-induced barrier lowering. Thus V_{th} is re-written with variables V_{th} and *Beta* as:

$$V_{th} = V_{constant} - B_{eta} V_{ds}. \quad \textbf{(Eq.3.3)}$$

Table3.1 shows the values of $\boldsymbol{V_{constant}}$ and $\boldsymbol{B_{eta}}$ of M1 and M2 used in the SCM (Fig. 2(a)) for different breakdown voltages. The observed values reveal that the different breakdown of transistors voltages have different $\boldsymbol{V_{th}}$ s even for the same process. Also $\boldsymbol{V_{th}}$ cannot be approximately regarded as the same value among different transistors. For example, in the condition of 1.2V supply voltage, 1.2V Tr./90nm simulation results demonstrate that $Vds1$ is 0.5V and $Vds2$ is about 0.026V, therefore $Vth1$ is 0.42V and $Vth2$ is about 0.396V.

$V_{th} = V_{constant} - B_{eta}V_{ds}$ 90nm process	1.2V Tr.		3.3V Tr.	
	M1	M2	M1	M2
$V_{constant}$(V)	0.38	0.38	0.43	0.43
B_{eta}	-0.08	-0.61	-0.10	-0.83
$V_{th} = V_{constant} - B_{eta}V_{ds}$ 65nm process	1.8V Tr.			
	M1	M2		
$V_{constant}$(V)	0.36	0.36		
B_{eta}	-0.07	-0.42		

Table3.1 $\boldsymbol{V_{th}}$ dependency on $\boldsymbol{V_{ds}}$

According to the re-definition of $\boldsymbol{V_{th}}$ in Eq. (3.3), the currents flowing through M8 and M9 are expressed as:

$$\boldsymbol{I_{M8} = S_{M8}I_t exp(\frac{V_{gs}-V_{th8}}{nV_T})}, \quad \textbf{(Eq.3.4)}$$

$$\boldsymbol{I_{M9} = S_{M9}I_t exp(\frac{V_{gs}-V_{th9}}{nV_T})}. \quad \textbf{(Eq.3.5)}$$

Where $\boldsymbol{V_{gs}}$ is the drain to source voltage. We also obtain that

$$\boldsymbol{I_{M8} = \frac{S_{M6}}{S_{M7}} I_{M9}}. \quad \textbf{(Eq.3.6)}$$

Using Eq. (3.4), (3.5) and (3.6), the drain to source voltage of transistor M2 is

$$V_{ds2} = P_2 + B_{eta8}V_{ds8} - B_{eta9}V_{ds9}. \quad \textbf{(Eq.3.7)}$$

TransistorM1 is in the saturation region, transistorsM2 and M3 are in the triode region. Thus, we have:

$$I_{M1} = \frac{1}{2}k'S_{M1}(V_{ref} - V_{net1} - V_{t1})^2, \quad \textbf{(Eq.3.8)}$$

$$I_{M2} = k'S_{M2}(V_{ref} - V_{net2} - V_{t2} - \frac{V_{ds2}}{2})V_{ds2}, \quad \textbf{(Eq.3.9)}$$

$$I_{M3} = k'S_{M3}(V_{ref} - V_{t3} - \frac{V_{net2}}{2})V_{net2}. \quad \textbf{(Eq.3.10)}$$

By using Eq. (3.8) and (3.9), we can obtain that

$$V_X = V_{ref} - V_{net1} - V_{t1} \quad \textbf{(Eq.3.11)}$$
$$= \frac{P_1V_{ds2}}{2} + \sqrt{\frac{3P_1V_{ds2}^2}{4} + P_1V_{ds2}(V_{t1} - V_{t2})}.$$

Then substituting Eq. (3.11) into Eq. (3.8), the reference current I_{ref} is calculated as:

$$I_{ref} = \frac{1}{2}k'S_{M1}(P_2 + V_{ds8}B_{eta8} - V_{ds9}B_{eta9})[\frac{P_1}{2} + \sqrt{\frac{3P_1}{4} + \frac{P_1B_{eta1}(P_2+V_{ds8}B_{eta8}-V_{ds9}B_{eta9}-V_{ds1})}{P_2+V_{ds8}B_{eta8}-V_{ds9}B_{eta9}}}]. \quad \textbf{(Eq.3.12)}$$

V_{ds2} is a function subject to V_{ds8} and V_{ds9}. Compared to Eq. (3.1), Eq. (3.12) considers the huge channel length transistors M1 and M2. They cause

larger slopes at output reference current when sweeping the power supply voltage. Furthermore, $\boldsymbol{V_{ds8}B_{eta8} - V_{ds9}B_{eta9} \ll V_{ds1}}$. Therefore, Eq. (3.12) also predicts that a large voltage drop at M1 is the dominant component to impact the stability of the reference current.

Considering the re-definition of $\boldsymbol{V_{th}}$, we can see that the dominant component to impact the stability of the reference voltage are both M1 and M2 as described in the following:

$$\mathbf{V}_{\boldsymbol{ref}} = \boldsymbol{V_{constant2} + P_3 + (P_2 + P_6)}$$

$$\times \sqrt{\frac{\boldsymbol{S_{M2}}}{\boldsymbol{S_{M3}}}\boldsymbol{(P_5 + P_3)(2P_4 - 1) + P_4^2}}. \qquad \textbf{(Eq.3.13)}$$

Where $\boldsymbol{P_3 = B_{eta1}(P_2 + V_{ds8}B_{eta8} - V_{ds9}B_{eta9} - V_{ds1})}$,

$$\boldsymbol{P_4 = 1 + P_1 + \sqrt{1 + P_1^2}},$$

$\boldsymbol{P_5 = \frac{S_{M5}+S_{M6}+S_{M7}}{S_{M5}+S_{M6}}}$ and

$$\boldsymbol{P_6 = V_{ds8}B_{eta8} - V_{ds9}B_{eta9}}.$$

One observation from Eq. (3.12) and Eq. (3.13) is that, for migrating reference circuit, the different $\boldsymbol{V_{th}}$s in various manufacturing processes demand changing of the transistor's geometrical size ratios *S M*. This conclusion verifies that the transistor's geometrical size ratios from the original process should be tuned once again to the target process in order to minimize the slope factor.

3.3 Analyzing SCM and SBC based on ACM model

Since we plan to migrate the reference circuit between two processes, a rapid retargeting methodology is required to satisfy the growing of time-to-market demands. Suppose we apply the derivations in Section 3.2 to implement an automation computer aided design, the known parameters include the target specifications $\boldsymbol{I_{ref}}$ and $\boldsymbol{V_{ref}}$, the original ratios of geometric size and the technological parameters of each process. Except the geometric size ratio of M10, all other geometric size ratios are associated to the values of reference current and reference voltage. Even the influence of M6 and M5 is very small, the geometric size of M6 and M5 can be directly obtained by scaling factors. Solving the unknowns $S\,M1$, $S\,M2$, $S\,M3$, $S\,M8$ and $S\,M9$ in Eq. (3.12) and Eq. (3.13) is a tough work. It requires a lot of assumptions and empirical predictions. Hence the above methods are not applicable to automation computer aided design.

Here, an advanced compact MOSFET (ACM) model is used for introducing the concept of inversion level. ACM also enables us to develop a ratio-based design with a good portability in the process migration.

The ACM model employs the concept of inversion level [50]. It can describe MOSFET behaviors for all operating regions. According to the ACM model, the drain current can be split into forward ($\boldsymbol{I_F}$) and reverse ($\boldsymbol{I_R}$) currents. We notice that only when a transistor operates in the saturation region, the reverse current can be neglected. We present several basic formulas in the following.

$$\boldsymbol{I_D = I_F - I_R = I_S(i_f - i_r)}, \qquad \textbf{(Eq.3.14)}$$

$$\boldsymbol{I_S = I_{SQ}\left(\frac{W}{L}\right) = I_{SQ}(S)}, \qquad \textbf{(Eq.3.15)}$$

$$I_{SQ} = \mu C'_{OX}\eta\frac{V_T^2}{2}. \qquad \textbf{(Eq.3.16)}$$

Where i_f and i_r are the forward and reverse inversion coefficients, respectively. I_{SQ} is the sheet normalization current. μ is the channel effective mobility. Note that μ is slightly dependent on the gate voltage V_G in actual operation. C'_{OX} is the gate capacitance per unit area.

The relationship between the current and the voltage is described as:

$$\frac{V_P - V_{S(D)}}{V_T} = \sqrt{1 + i_{f(r)}} - 2 + ln(\sqrt{1 + i_{f(r)}} - 1). \qquad \textbf{(Eq.3.17)}$$

Where $V_P \approx (V_G - V_{T0})/\eta$ is the pinch-off voltage and V_{T0} is the zero bias threshold voltage. Herein, Eq. (3.17) can be represented as $V_{S(D)} = V_P - V_T \times F_{(i_{fn})}$.

Eq. (3.12) and Eq. (3.13) suggest that the reference circuit should be tuned by changing the geometrical sizes of M1, M2, M3, M8 and M9.

First of all, following the ACM model, we reveal that the ratio $\frac{V_{ds2}}{V_{ds1}+V_{ds2}}$ can be regarded as a constant proportional coefficient. The SCM voltage proportional coefficient of Q is defined as:

$$Q = \frac{V_{ds2}}{V_{ds1}+V_{ds2}}. \qquad \textbf{(Eq.3.18)}$$

The ACM model applying to SCM gives a key portion that M1 is working in the saturation region and M2 is in the deep triode region. Applying Eq. (3.17) to both source terminals of M1 and M2, the following is obtained.

$$V_{s1} - V_{s2} = V_T[F_{(i_{f2})} - F_{(i_{f1})}]. \qquad \textbf{(Eq.3.19)}$$

Similarly, applying Eq. (3.17) to the drain and source terminals of M2:

$$V_{ds2} = V_T[F_{(i_{f2})} - F_{(i_{f1})}]. \qquad \textbf{(Eq.3.20)}$$

As shown in Fig.3.2 (a), a main bias current is given as I_d, and an optional current is as NI_d. Both M1 and M2 share the same gate voltage and M1 is tied to the drain of M2, that is, i_{f1}=i_{r2}. M1 is saturated, so $i_{f1} \gg i_{r1}$. Accordingly,

$$I_{d1} = I_d = S_1 I_{SQ} i_{f1},$$

$$I_{d2} = (N+1)I_d = S_2 I_{SQ}(i_{f2} - i_{f1}). \qquad \textbf{(Eq.3.21)}$$

This results in i_{f2}=$G_1 i_{f1}$, where

$$G_1 = [1 + (N+1)\frac{S_1}{S_2}]. \qquad \textbf{(Eq.3.22)}$$

In the same situation of M1 and M2, M2 and M3 share the same gate voltage andM2 is tied to the drain ofM3, that is, i_{f2}=i_{r3}. Suppose the current flows through M9 is M times of I_d, then I_{d3} can be expressed as:

$$I_{d3} = (N+M+1)I_d = S_3 I_{SQ}\left(i_{f3} - i_{f2}\right). \qquad \textbf{(Eq.3.23)}$$

Thus i_{f3}=$G_2 i_{f1}$, where

$$G_2 = [G_1 + (N+M+1)\frac{S_1}{S_3}]. \qquad \textbf{(Eq.3.24)}$$

The application of Eq. (3.17), Eq. (3.22) and Eq. (3.24) to node V_{S2} for M2 results in:

$$\boldsymbol{V_{S2} = V_{S3} + V_T[F(i_{f3}) - F(i_{f2})]}$$

$$\boldsymbol{= \left[\sqrt{1 + G_2 i_{f1}} + ln\left(\sqrt{1 + G_2 i_{f1}} - 1\right)\right]}$$

$$\boldsymbol{-\left[\sqrt{1 + G_1 i_{f1}} + ln\left(\sqrt{1 + G_1 i_{f1}} - 1\right)\right].} \qquad \textbf{(Eq.3.25)}$$

According to Eq. (3.20) and Eq. (3.22), we obtain the following equations:

$$\frac{V_{ds2}}{V_T} = F\left(G i_{f1}\right) - F\left(i_{f1}\right)$$

$$\boldsymbol{= \left[\sqrt{1 + G i_{f1}} + ln\left(\sqrt{1 + G i_{f1}} - 1\right)\right]}$$

$$\boldsymbol{-\left[\sqrt{1 + i_{f1}} + ln\left(\sqrt{1 + i_{f1}} - 1\right)\right].} \qquad \textbf{(Eq.3.26)}$$

Combining Eq. (3.25) and Eq. (3.26), we calculate V_{s1} as:

$$\boldsymbol{V_{S1} = V_{S2} + V_{ds2}}$$

$$\boldsymbol{= \left[\sqrt{1 + G_2 i_{f1}} + ln\left(\sqrt{1 + G_2 i_{f1}} - 1\right)\right]}$$

$$\boldsymbol{-\left[\sqrt{1 + i_{f1}} + ln\left(\sqrt{1 + i_{f1}} - 1\right)\right].} \qquad \textbf{(Eq.3.27)}$$

Noticed that $\boldsymbol{V_{d1}} = \boldsymbol{V_{g1}}$, the drain to source voltage of M1 is obtained by solving the Eq. (3.17) and Eq. (3.27) simultaneously.

$$\boldsymbol{V_{ds1} = V_{gs1} = nV_T\left[\sqrt{1 + i_{f1}} - 2 + \ln\left(\sqrt{1 + i_{f1}} - 1\right)\right]}$$

$$\boldsymbol{+V_{T0} + (n-1)V_{S1}.} \qquad \textbf{(Eq.3.28)}$$

Eq. (3.21) suggests that $\boldsymbol{i_{f1}}$ is proportional to $\boldsymbol{I_d}$. That is, $\boldsymbol{i_{f1}}$ is a constant since $\boldsymbol{I_d}$ (i.e. $=\boldsymbol{I_{ref}}$) is specified by the requirement. Reference voltage is given as $\boldsymbol{V_{ref} = V_{ds1}+V_{s1}}$. We can observe the sensitivity of the reference current to the reference voltage subject to $\boldsymbol{i_{f1}}$. Low inversion levels of $\boldsymbol{i_{f1}}$ made the sensitivity extremely high. For instance, $\boldsymbol{i_{f1} \leq 6}$ in 1.2V Tr./90nm process causes the reference current changing more than 20%/*V*. In the design case, M1 is operated in saturation region with $\boldsymbol{i_{f1} = 11.5}$. Note that the detailed analysis about $\boldsymbol{i_{f1}}$ and $\boldsymbol{i_{f2}}$ can be found in [41] and [50].

As a function of $\boldsymbol{i_{f1}}$, the SCM voltage proportional coefficient of Q can be regarded as a constant proportional coefficient.

All the above derivations are based on the conditions that M1, M2 and M3 have the same threshold voltage of $\boldsymbol{V_{th}}$. In our simulations, the ratio Q is from 4% to 6% by considering different $\boldsymbol{V_{th}}$s (in the case that *N* = 1, *M* = 1). We suggest *Q* = 5% for the coarse tuning.

Eq. (3.19) to Eq. (3.28) are certification procedures to set forth that Q can be expressed as a constant proportional coefficient. Based on the deduced property of Q, we adopt a coarse tuning parameter of K to repeat sweeping the drain to source voltage of M2. As previously mentioned in Eq. (3.7), $\boldsymbol{V_{ds2}}$ is approximately equal to $\boldsymbol{P_2 = nV_T ln(\frac{S_{M6}S_{M8}}{S_{M7}S_{M9}})}$. In this design of SBC circuit, M6 and M7 have the same channel size. Then we set K as:

$$\boldsymbol{K = \frac{S_{M8}}{S_{M9}}}. \qquad \textbf{(Eq.3.29)}$$

Applying the scaling factors in [47] and [48], we obtain the minimum channel lengths of M8 and M9. As a result, *K* induces the channel widths of

M8 and M9. Our simulation results showed that the empirical range of K was from 2 to 6. (Noticed that the range is established also in the situation of $N = 1, M = 1$.)

Next, we clarify that the range of K is enough for migration of the ratio Q mentioned above. Hereinafter, we make use of the actual manufactured 1.2V transistor/90nm as an example. The sub-threshold slope factor $n \simeq 0.94$. $\boldsymbol{V_{ds1}} + \boldsymbol{V_{ds2}} \simeq 0.34V$ when $K = 2$. $\boldsymbol{V_{ds1}} + \boldsymbol{V_{ds2}} \simeq 0.88V$ when $K = 6$.

The sum of voltages of MOSFETs along a path from the power supply to the ground is bounded by the minimum supply voltage. The minimum supply voltage is obtained as follows:

$$\boldsymbol{VDD_{min} > V_{dsM4} + V_{gsM1} + V_{dsM2} + V_{dsM3}}$$
$$\text{or } \boldsymbol{V_{gsM6} + V_{dsM8} + V_{dsM2} + V_{dsM3}}. \quad \textbf{(Eq.3.30)}$$

Hence, since the threshold voltage of NMOS transistor is 0.42V, the supply voltage can be as low as 0.65V. The maximum sum 0.88V of $\boldsymbol{V_{ds1}}$ and $\boldsymbol{V_{ds2}}$ already exceeds the minimum supply voltage.

In other simulation results, the range of K is [2, 6], the range of $\boldsymbol{V_{ds1}}$ + $\boldsymbol{V_{ds2}}$ is [0.47V, 1.21V] ($n \simeq 1.45$) in the usage of 3.3V Tr./90nm. The minimum supply voltage is 0.8V. In another case, the range of $\boldsymbol{V_{ds1}} + \boldsymbol{V_{ds2}}$ is [0.38V, 1.05V] ($n \simeq 1.2$) by utilizing 1.8V Tr./65nm, and the minimum supply voltage is 0.68V. These results convince us that our range of K gives a good tuning range.

3.4 Retargeting flow with tuning procedure

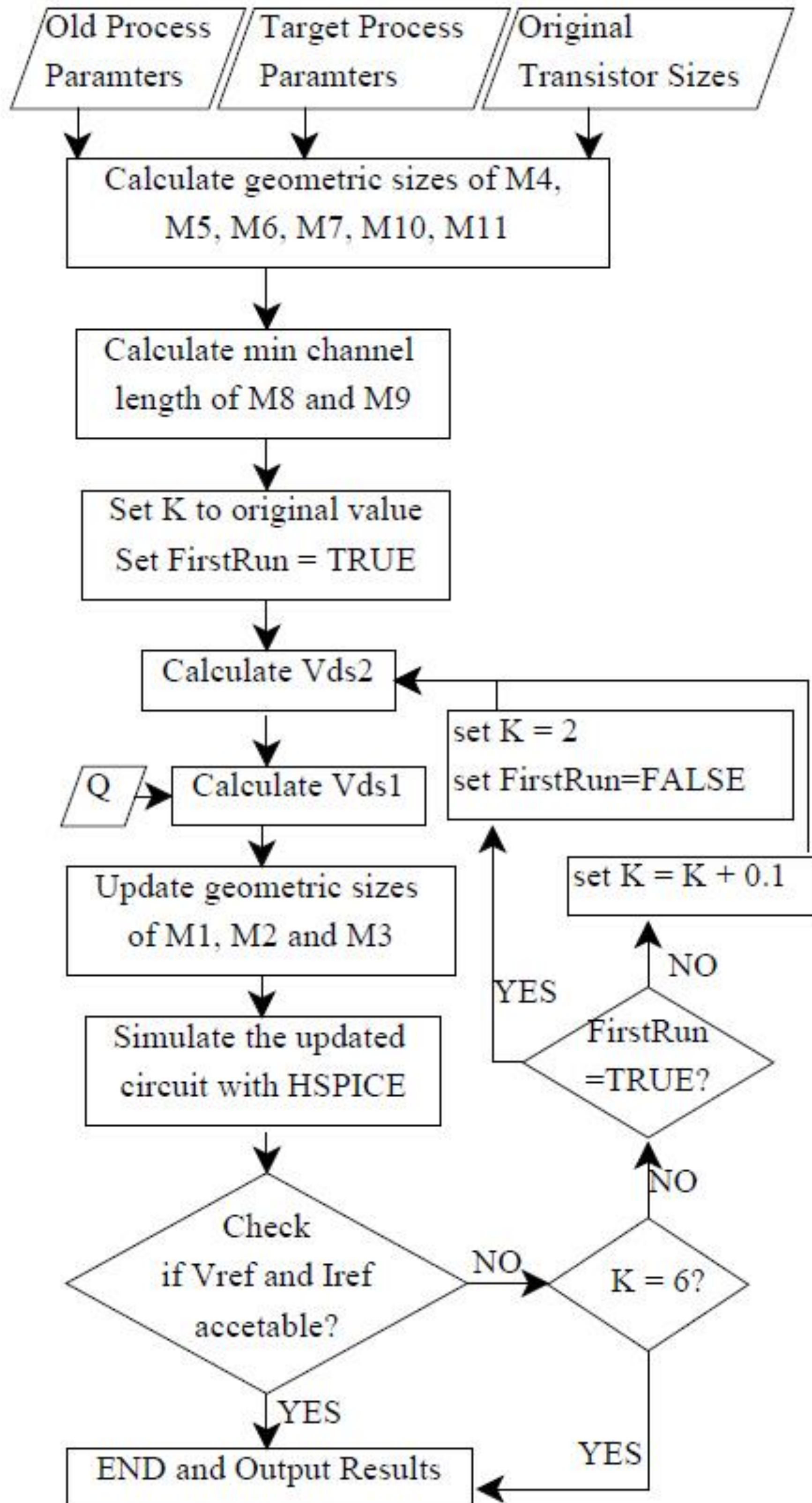


Fig.3.3 Proposed tuning process for parameters determination

The tuning steps for the reference circuit are summarized in Fig.3.3. We explain the flow chart as follows;

① The original and target process parameters, original geometric sizes are read in as input information.

② Retargeting the channel sizes of M4, M5, M6, M7, M10 and M11 by using scaling factors. Then, calculate the minimum channel length of M8 and M9. The initial *K* value is set to the same as the original process. Adjust *K* coarsely to obtain V_{ds2} (here, we only use $N = 1$, $M = 1$). The simulation results suggest that the range of *K* is from 2 to 6.

③ Calculate V_{ds2} by the equation $V_{ds2} = nV_T ln(K)$. The empirical value of V_{ds1} within the range of the SCM voltage proportional coefficient Q is from 94% to 96%. (Here, in the case of $N = 1$, $M = 1$, we suggest adopting 95% as the coarse tuning.) Hence, the aspect ratio S 1, S 2 and S 3 can be calculated from Eq. (3.21) and Eq.(3.23).

④ If the simulation shows good results, then this procedure is terminated. If the simulation results do not satisfy the specifications, then *K* is repeated to sweep in the range of [2-6] except the original ratio. In total, the methodology is executed at most 40 times. Finally, the optimal solution is picked up by designers.

The procedure is for coarse tuning. Firstly, the original *K* is supplied to observe the circuit performance. If original *K* setting does not meet the requirement, then *K* is performed in a range of [2, $V_{initial}$)($V_{initial}$, 6]. Thus this procedure is repeated at most 39 times. Finally, the optimal solution is picked up by designers. The duration for the methodology's automatic calculation and using HSPICE to simulate this reference circuit once does

not exceed one second. Even if 40 times sweeping are needed, the total time is less than one minute, which is much faster than manual design.

3.5 Leakage current compensation by start-up circuit

A conventional start-up circuit is shown in Fig.3.4 (a) (named Start-up (A)). As a comparison of Start-up (A), we prepare another start-up circuit shown in Fig.3.4 (b) (named Start-up (B)). Startup (B) is designed without resistors. The start-up circuit enables the reference circuit to satisfy the stability over the wide range of the supply voltage.

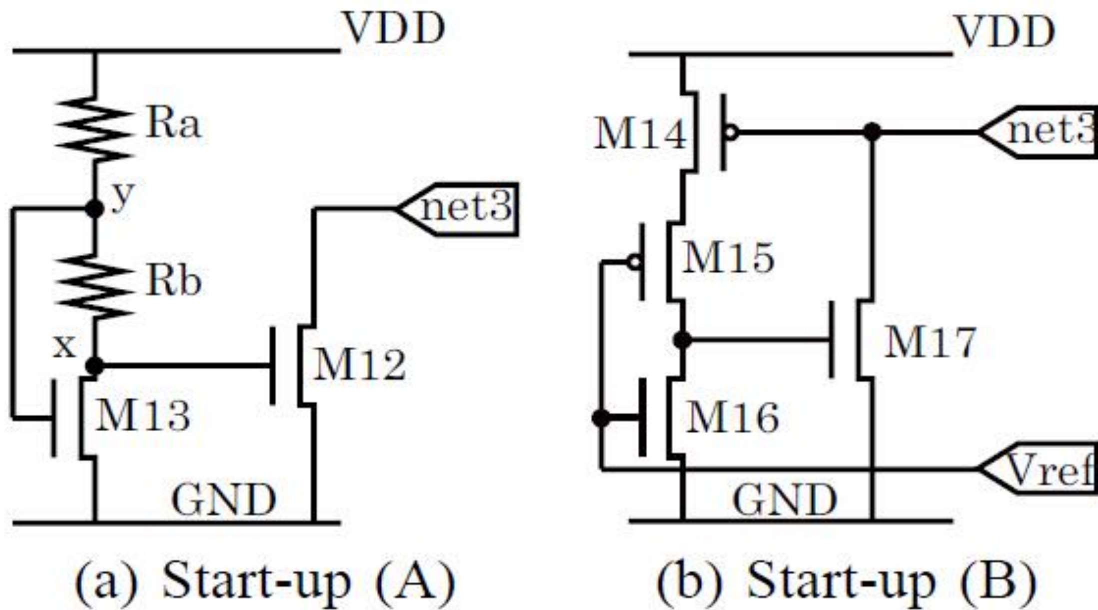


Fig.3.4 Leakage current compensation Start-up circuit

For Start-up (A), the values of parameters must be given in the Eq. (3.31) so that M12 turns off.

$$V_{DD} - I_{DSM13}(R_a + R_b) = V_X = V_{gs2} \leq V_{th12} \qquad \textbf{(Eq.3.31)}$$

$\boldsymbol{I_{DSM12}}$ supplies about $\mathbf{10^{-12}}$ ampere leakage current to the net3 in the SBC circuit. The resistors $\boldsymbol{R_a}$ and $\boldsymbol{R_b}$ play a role to increase the voltage partially. Suppose there are two voltage points $\boldsymbol{V_{DD1}}$ and $\boldsymbol{V_{DD2}}$ in the range of supply voltage when M12 is on, and $\boldsymbol{V_{DD2}} > \boldsymbol{V_{DD1}}$. Let $\boldsymbol{I_D}$ of M13 at $\boldsymbol{V_{DD1}}$ and $\boldsymbol{V_{DD2}}$ to be $\boldsymbol{I_1}$ and $\boldsymbol{I_2}$, respectively. $\boldsymbol{V_{DD2}}$-$\boldsymbol{V_{DD1}}$ < $\boldsymbol{I_2(R_a + R_b)}$-$\boldsymbol{I_1(R_a + R_b)}$. Also the gate-source voltage in M12 is $\boldsymbol{V_{gs1}}$ and $\boldsymbol{V_{gs2}}$, $\boldsymbol{V_{x2}} = \boldsymbol{V_{gs2}} < \boldsymbol{V_{x1}} = \boldsymbol{V_{gs1}}$. Hence, $\boldsymbol{V_{DD}}$ increases, and then $\boldsymbol{V_{gs}}$ of M12 decreases. However, from Eq. (3.31), $\boldsymbol{V_{gs}}$ of M12 increases with the value of $\boldsymbol{V_{DD}}$. Thus the start-up circuit can make sure transistor of M12 working as soon as possible.

3.6 LDE-aware LCR be utilized in the reference circuit

In [42], it mentions that M5 is added to compensate those leakage currents to make the drain current of M8 and M9 equal. And as proved in the simulation results, the temperature performance can be improved by the reverse bias diode M5. Actually, the reason is that M8 and M9 has different channel width. As shown in Fig.2.1 (a), P^+ sidewall cause different junction leakage associate with different channel width. That is why $\boldsymbol{I_{M8}}$ could not strictly follow the Eq. (3.6), to be demission ratio of M6 and M7. Even if the leakage is in dozens of pico-ammeter level, it is too sensitive to the nano-ammeter reference current.

Instead of reverse bias diode M5, it is a clever application to use LDE-aware LCR in this reference circuit. In the simulations, tuning M8 has wider diffusion size can limited the leakage in transistor M8. To make sure junction leakage from P^+ sidewall as equal as possible, the simulations can show

the same effect on leakage compensation of M5. Therefore, the application of LDE-aware LCR can reduce the complexity of the reference of circuit.

3.7 Temperature Compensation

Since the effective gain factor ($\boldsymbol{n^2 k'}$) varies with the temperature in the same way as the mobility, it can be expressed as [51]

$$(\boldsymbol{n^2 k'}) = (\boldsymbol{n^2 k'})_0 (\frac{T_0}{T})^m. \qquad \textbf{(Eq.3.32)}$$

Thus, the reference current $\boldsymbol{I_{ref}}$ can be expressed as

$$\boldsymbol{I_{ref}} = (\boldsymbol{n^2 k'})_0 V_{T_0}^2 K_{eff} (\frac{T_0}{T})^{2-m}. \qquad \textbf{(Eq.3.33)}$$

By differentiating Eq. (3.33) with respect to the temperature, we can obtain that

$$\frac{\partial I_{ref}}{\partial T} = (2 - \mathrm{m})(\boldsymbol{n^2 k'})_0 V_{T_0}^2 K_{eff} (\frac{T_0}{T})^{1-m}. \qquad \textbf{(Eq.3.34)}$$

Where $(\boldsymbol{n^2 k'})_0$ and $\boldsymbol{V_{T_0}}$ are defined as the parameters at room temperature $\boldsymbol{T_0}$. And *m* is between 1.5 and 2, which makes the exponent (2 − *m*) very small. Thus the reference current $\boldsymbol{I_{ref}}$ is almost insensitive to temperature.

The threshold voltage of NMOS transistor linearly decreases with temperature and can be expressed as the first-order approximation as

$$\boldsymbol{V_t(T) = V_t(T_0) - K_{TN}(T - T_0)}, \qquad \textbf{(Eq.3.35)}$$

where $\boldsymbol{K_{TN}}$ is the threshold voltage temperature coefficient. Then the reference voltage $\boldsymbol{V_{ref}}$ can be expressed as

$$V_{ref} = V_t(T_0) - K_{TN}(T - T_0) +$$

$$nV_{T_0}\frac{T}{T_0}ln(P_1)\sqrt{\frac{S_{M2}}{S_{M3}}P_5(2P_4 - 1) + (P_4)^2}. \quad \textbf{(Eq.3.36)}$$

By differentiating Eq. (3.36) with respect to temperature, we can obtain that

$$\frac{\partial V_{ref}}{\partial T} = -K_{TN} + n\frac{V_{T_0}}{T_0}ln(P_1)\sqrt{\frac{S_{M2}}{S_{M3}}P_5(2P_4 - 1) + (P_4)^2}. \quad \textbf{(Eq.3.37)}$$

To choose the transistor sizes, the effect of temperature for V_{ref} can be cancelled, which means that $\frac{\partial V_{ref}}{\partial T} = 0$ and

$$ln(P_1)\sqrt{\frac{S_{M2}}{S_{M3}}P_5(2P_4 - 1) + (P_4)^2} = K_{TN}/(\frac{nV_{T_0}}{T_0}). \quad \textbf{(Eq.3.38)}$$

Substituting Eq. (3.38) into Eq. (3.36), we can get the reference voltage when its temperature coefficient is zero as

$$V_{ref} = V_t(T_0) + K_{TN}T_0 \quad \textbf{(Eq.3.39)}$$

3.8 Results

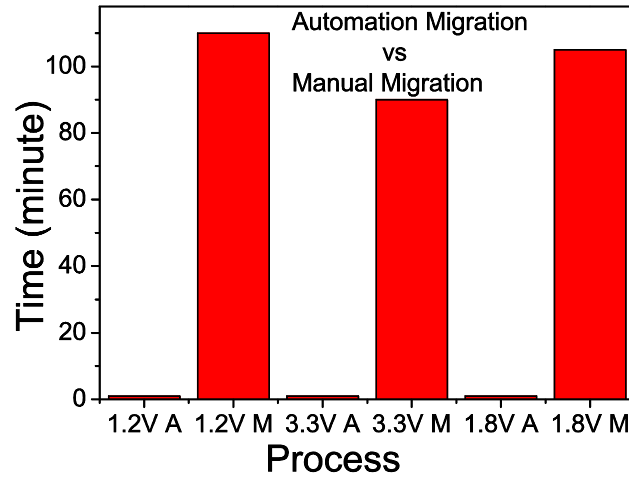
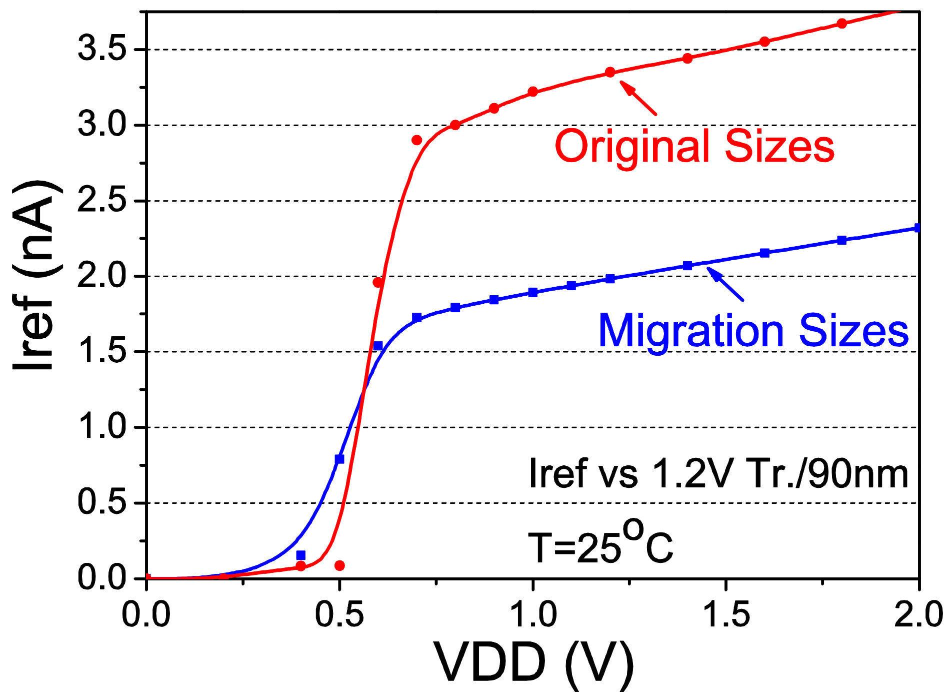
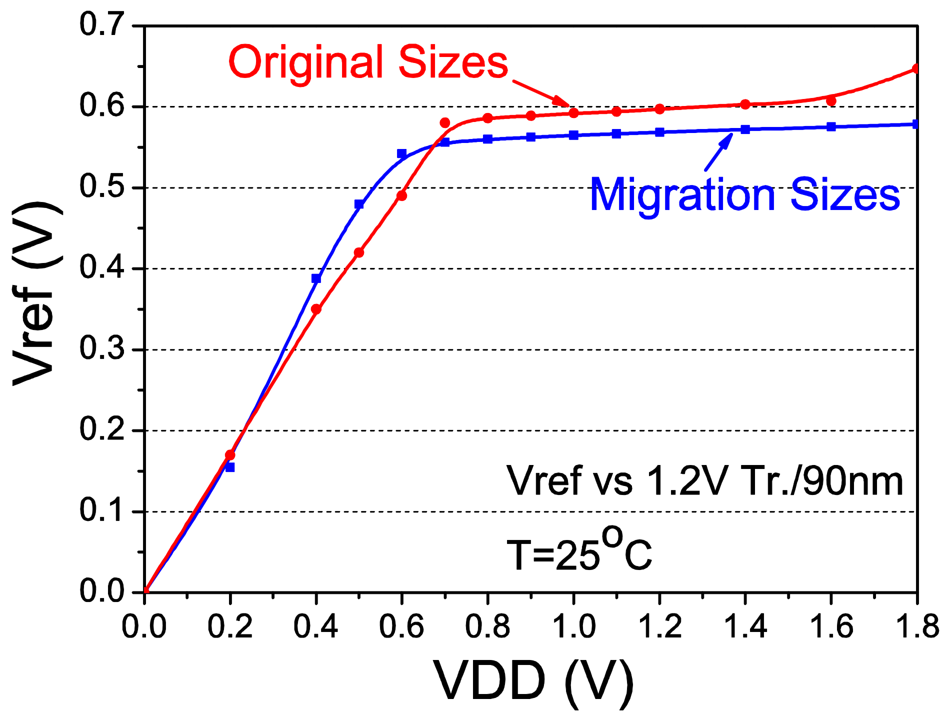


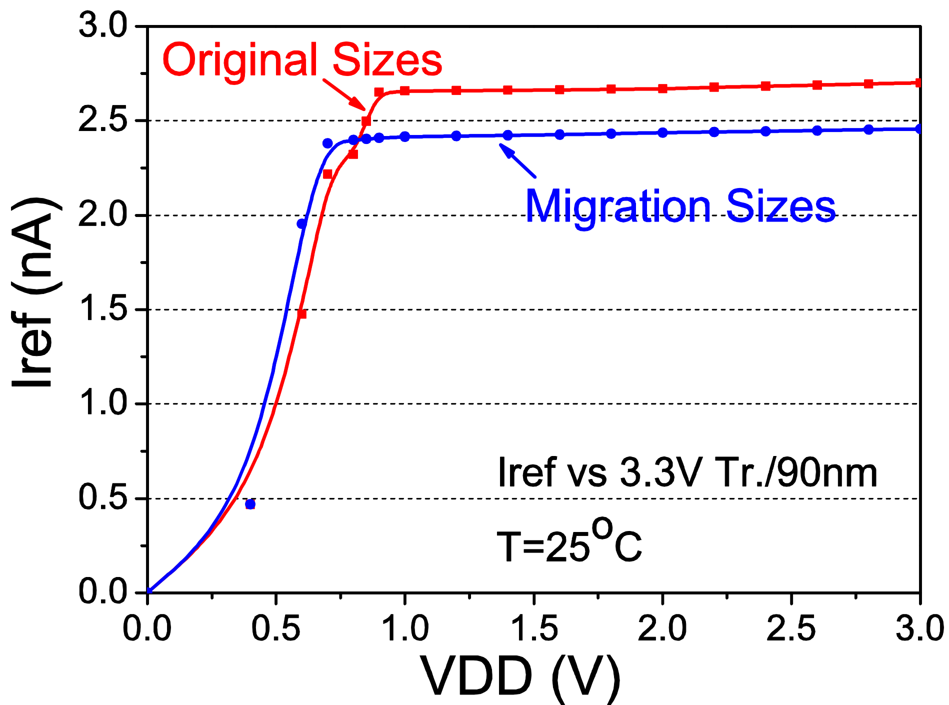
Fig.3.5 A comparison of spending time between automatic migration and manual migration



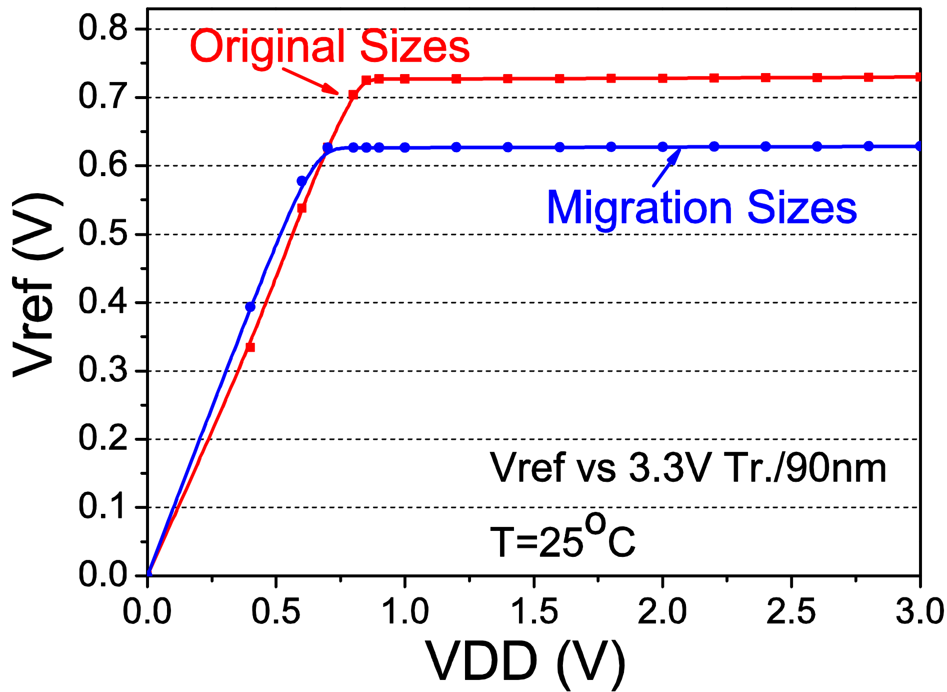
(a) 1.2V Tr./90nm I_{ref} vs. V_{DD}



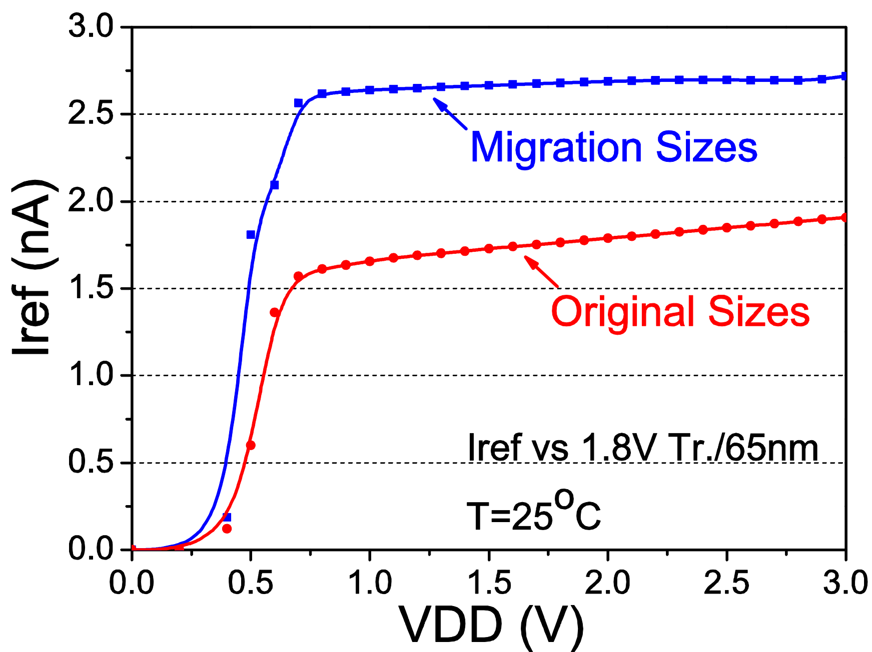
(b) 1.2V Tr./90nm $\boldsymbol{V_{ref}}$ vs. $\boldsymbol{V_{DD}}$



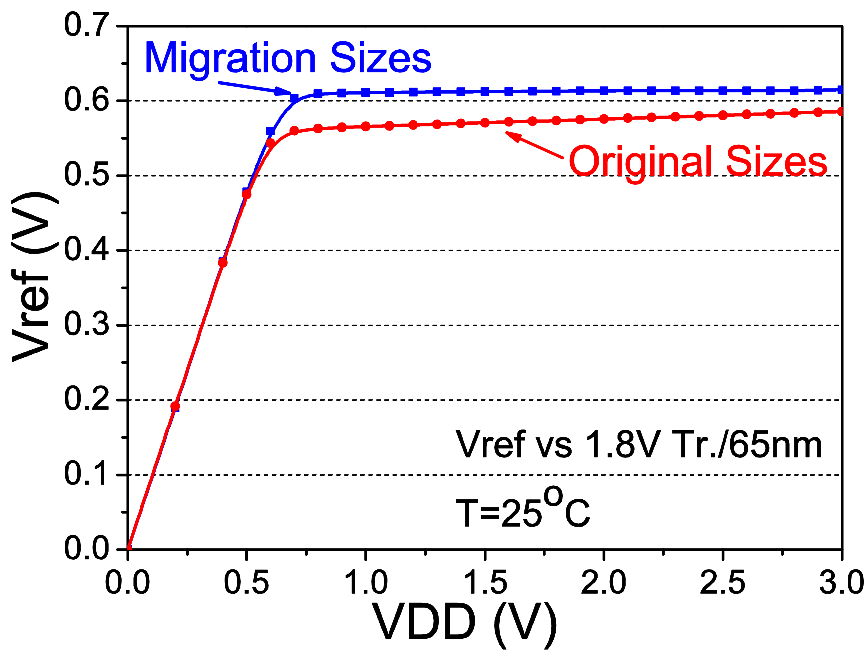
(c) 3.3V Tr./90nm $\boldsymbol{I_{ref}}$ vs. $\boldsymbol{V_{DD}}$



(d) 3.3V Tr./90nm $\boldsymbol{V_{ref}}$ vs. $\boldsymbol{V_{DD}}$

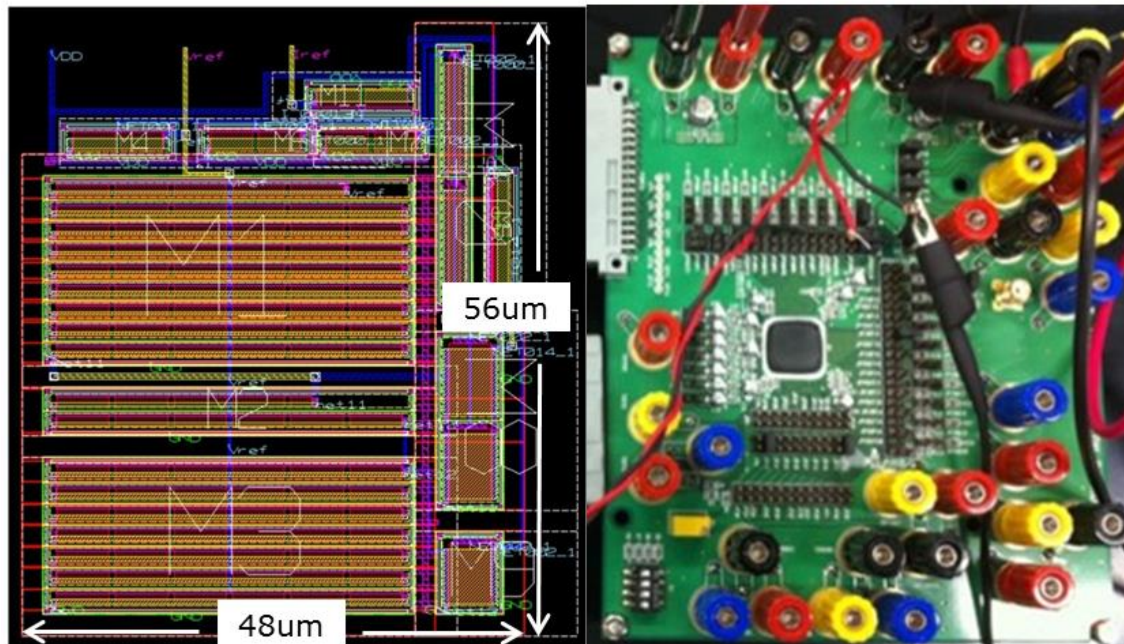


(e) 1.8V Tr./65nm $\boldsymbol{I_{ref}}$ vs. $\boldsymbol{V_{DD}}$



(f) 1.8V Tr./65nm V_{ref} vs. V_{DD}

Fig.3.6 A comparison between original simulation results and migration simulation results of reference voltage (V_{ref}) and reference current (I_{ref}) at room temperature



(a) Layout of the Reference Circuit

(b) Headstage board

Fig.3.7 The layout of the reference circuit and the head-stage board

The reference circuit introduced previously (Fig.3.1) was originally designed for 180nm technology with a maximum supply voltage of 1.8V. In this work, we attempt to retarget it to other foundries, using 90nm technology with 1.2V, 3.3V transistors and 65nm technology with 1.8V transistors to verify our methodology. We design the reference circuit following the specifications of $\boldsymbol{I_{ref}}= 2nA$ and $\boldsymbol{V_{ref}}= 0.6V$.

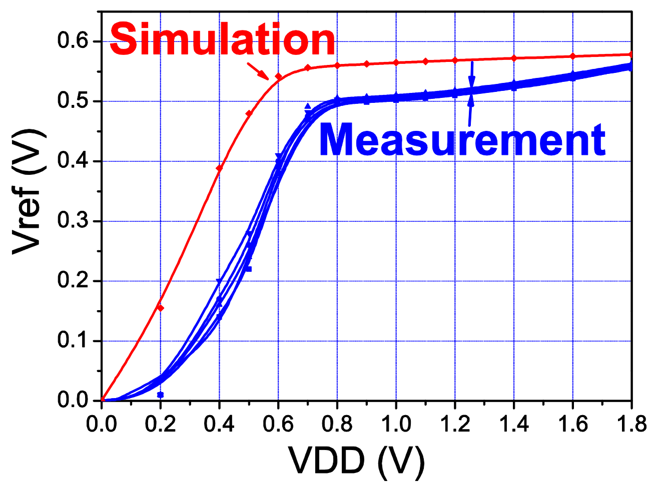
The post layout simulation results and measurement results using BSIM level 54 model are given in Table3.2. The purpose of listing two types of simulation to each category of transistor is to verify the importance of retargeting transistor's geometric sizes in migration. Original simulation results proved the original process transistor's geometric sizes to the target process directly. Migration simulation results are carried out with our methodology to obtain a set of new transistor's geometric sizes. The

parameter	1.2V Tr. 90nm Pro.			3.3V Tr. 90nm Pro.		1.8V 65nm Pro.		unit
	k=2.5	k=2.4		k=2.5	k=3.3	k=2.5	k=3.2	
	Original Sim.	Migration Sim.	Mea.	Original Sim.	Migration Sim.	Original Sim.	Migration Sim.	
VDDmin	0.68	0.65	0.65	0.82	0.73	0.7	0.68	V
Power(at1.1V)	3.52	1.55	1.1	2.9	2.66	1.85	2.9	nW
Vref	0.6	0.55	0.5	0.727	0.627	0.57	0.61	V
Vref sensitivity to VDD	2.0	1.92	2.04	0.101	0.08	0.193	0.09	%/V
Vref sensitivity to T	0.005	0.005	0.011	0.08	0.08	0.05	0.05	%/°C
Iref	3.2	1.5	1	2.66	2.42	1.68	2.7	nA
Iref sensitivity to VDD	12.07	10.56	12.8	2.84	2.42	6.91	3.11	%/V
Iref sensitivity to T	0.227	0.228	0.694	0.28	0.279	0.287	0.287	%/°C

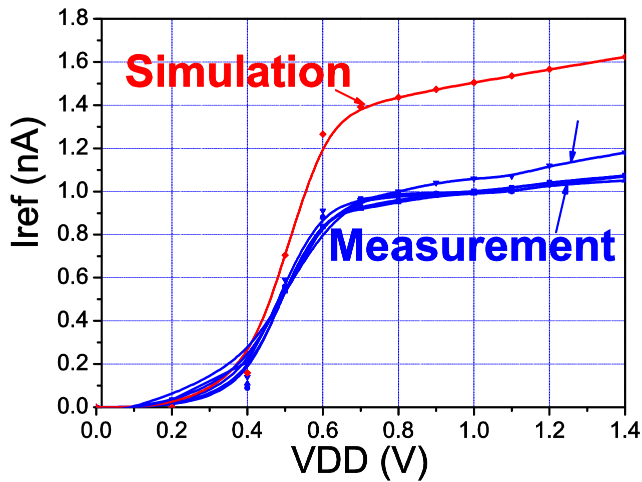
Table3.2 Summary of simulations and measurements

process	Original Process		Target Process					
Transistor	1.8V Tr./180nm		1.2V Tr./90nm		3.3V Tr./90nm		1.8V Tr./65nm	
	W[um]	L[um]	W[um]	L[um]	W[um]	L[um]	W[um]	L[um]
M1	0.8	260	0.3	265	1	300	0.5	410
M2	2	200	0.25	52	1	140	0.5	110
M3	0.8	340	0.3	210	1	100	0.5	100
M4,6,7,11	4	100	1	25	4	100	2	50
M5	20	100	1	5	2	10	0.5	2.5
M8	10	10	12	4	10	20	13	30
M9	4	10	5	4	3	20	4	30
M10	20	1	20	1	20	1	10	0.5

Table3.3 Transistor sizes for the reference circuit been migrated among different foundry's processes

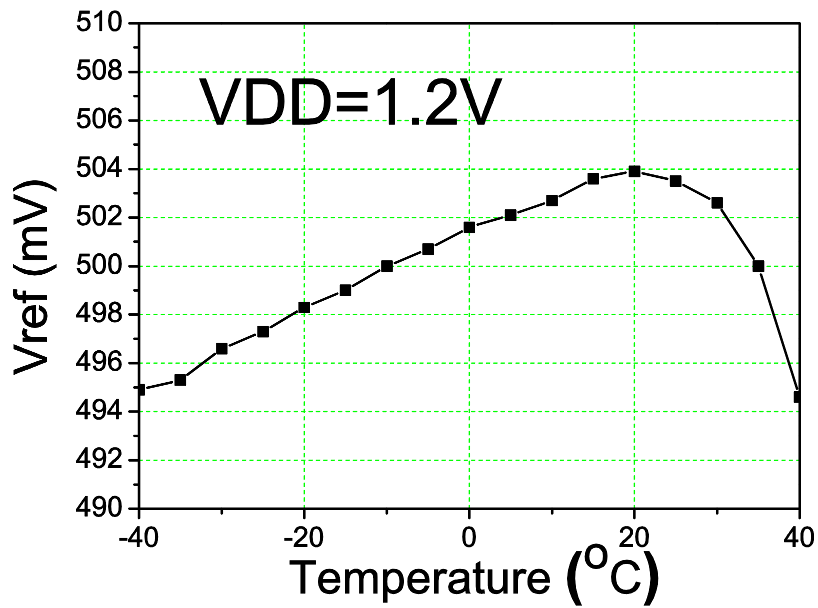


(a) Reference voltage

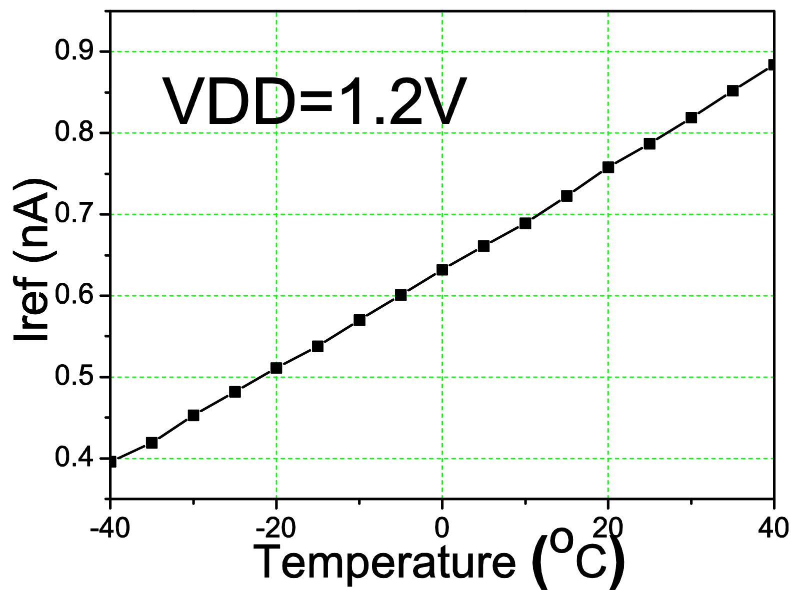


(b) Reference current

Fig.3.8 Comparison of reference voltage and reference current between simulation and measurement result at room temperature of 25℃



(a) Reference voltage



(b) Reference current

Fig.3.9 The measurement results of the $\boldsymbol{V_{ref}}$ and $\boldsymbol{I_{ref}}$ versus temperature

performance of reference circuit is improved and matched the requirements of specifications in the migration simulation results.

The channel sizes of each transistor are shown in Table3.3. The reference circuits of 3.3V/90nm and 1.8V/65nm with almost the same length parameters but different K value are able to attain the best performance. However, the 1.2V transistor resulted in smaller size according to the scaling factors.

Fig.3.5 compares the time spent between automatic migration and manual migration. By adopting automatic migration method, all the designs do not spend more than one minute, but, the time spent by manual migration for 1.2V/90nm, 3.3V/90nm and 1.8V/65nm designs are 110 minutes, 90 minutes and 105 minutes respectively. These results show that our proposed automatic migration method works much faster than the manual migration.

Although the manual migration time is quite different and heavily based on the experience of designers, to complete the migration in one minute is a tough work for most of them.

A comparison between original simulation results and migration simulation results for reference voltage ($\boldsymbol{V_{ref}}$) and reference current ($\boldsymbol{I_{ref}}$) at room temperature are given in Fig.3.6. The red line is simulated by designing reference circuit with original transistor's geometric sizes. The blue line is optimal performance chosen from 40 simulations.

The rough design suggests that the sensitivity of the reference current and the reference voltage to the supply voltage is relatively low and acceptable for most applications. Ratio of output voltage change to power supply voltage change (PSRR) is usually used as a criterion to the reference circuit, where, $PSRR = 20log\ (\Delta V_{power-supply} / \Delta V_{ref-output})$ dB. The resultant PSRR for 1.2V/90nm, 3.3V/90nm and 1.8V/65nm transistors are 31.8dB, 63dB and 62dB, respectively.

Fig.3.6 reveals that the slope degree becomes larger particularly when the low breakdown voltage transistors such as 1.2V is used. The reason is that low breakdown voltage MOSFETs technology has greater leakage current. For a nano-ampere level reference current, the leakage current ($\boldsymbol{I_{off}}$) should be smaller than the reference current. The reference circuit utilizing 1.2V/90nm transistors is fabricated. The NMOS type $\boldsymbol{I_{off}}$ is -0.256nA, PMOS type $\boldsymbol{I_{off}}$ is 0.295nA and the specification reference current is 1.5nA. For a high accuracy measurement of the reference current, we got rid of the ESD protection from the IO pads to prevent the leakage current from flowing. The layout of the reference circuit and the head-stage board are shown in Fig.3.7. The area is around $48 \times 56\mu m^2$.

The comparison between simulation and measurement results for reference voltage and reference current are shown in Fig.3.8. In total five samples of reference circuits measurement results are used to validate the design. The results show that at room temperature of 25℃, the circuit could operate normally at the supply voltage down to 0.6V. The average reference current is 1.07nA with a maximum deviation of ±7.1%at 1.2V supply. The average reference voltage is 0.522V with a maximum deviation of ±0.9% at 1.2V supply.

Fig.3.9 shows the measurement results of output reference voltage and reference current versus temperature from −40℃ to 40℃ when supply voltage VDD is 1.2V. Similar results could be found in previous work [41]. Over the temperature range of the measurement, the reference voltage achieved the highest voltage at 20℃. We also measured temperature range from40℃ to 90℃, however, for temperature close to 50℃ or higher, the current measured is much higher than expected. The PTAT (Proportional to Absolute Temperature) dependence of the $\boldsymbol{I_{ref}}$ and $\boldsymbol{V_{ref}}$ should be incorporated in future work of the retargeting methodology.

3.9 Conclusion Remarks

We presented a novel retargeting methodology for the advanced low power reference circuit. We clarified the dependency of $\boldsymbol{V_{th}}$ upon $\boldsymbol{V_{ds}}$, and introduced the automatic migration approach to conserve the stability of the reference current and voltage. Based on the analysis of tuning procedures, a rapid solution for determining the transistor's geometric size can be obtained using automation design in less than one minute. Besides, a start-up circuit is added to the reference circuit. Moreover, LDE-aware LCR is used to compensate the leakage currents, which could reduce the complexity of

reference circuit. The simulation and measurement results convinced us that this approach has the potential to extend to the other reference circuits that based on SCM and SBC.

References

[39] 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.

[40] H. J. Oguey and D. Aebischer, “CMOS current reference without resistance,” IEEE J. Solid-State Circuits, vol. SC-32, no. 7, pp. 1132-1135, Jul. 1997.

[41] E. M. Camacho-Galeano, C. Galup-Montoro, and M. C. Schneider, “A 2-nW 1.1-V self-biased current reference in CMOS technology,” IEEE Tran. on Circuits and Systems-II, Vol. 52, No. 2, pp. 61-65, Feb. 2005.

[42] Zhangcai Huang; Qin Luo; Inoue, Y., ”A CMOS Sub-l-V nanopower current and voltage reference with leakage compensation,” IEEE International Symposium on Circuits and Systems, pp.4069-4072, May. 2010

[43] R.Acosta, F.Silveira and P.Aguirre, Experiences on analog circuit technology migration and reuse,Proceedings of 15th Symposium on Integrated Circuits and Systems Design, pp.169-174, 2002

[44] Batterywala, S.H.; Bhattacharya, S.; Rajagopalan, S.; Ma, H.-K.T.; Shenoy, N.V., ”Cell Swapping Based Migration Methodology for Analog and Custom Layouts,” Quality Electronic Design,

2008. ISQED 2008. 9th International Symposium on , vol., no.,pp.450,455, 17-19 March 2008

[45] Tutelea, L.; Boldea, I., ”Induction motor electromagnetic design optimization: Hooke Jeeves method versus genetic algorithms,” Optimization of Electrical and Electronic Equipment (OPTIM), 2010 12th International Conference on , vol., no., pp.485,492, 20-22 May 2010

[46] Naumowicz, M.; Melosik, M.; Katarzynski, P., ”Technology migration of analogue CMOS circuits using hooke-jeeves algorithm and genetic algorithms in multi-core CPU systems,” Mixed Design of Integrated Circuits and Systems (MIXDES), 2013 Proceedings of the 20th International Conference , vol., no., pp.267,272, 20-22 June 2013

[47] A. Savio, L. Colalongo, M. Quarantelli, and Z.M. Kovacs-Vajna,”Automatic Scaling Procedures for Analog Design Reuse,” IEEE Transactions on Circuits and Systems, vol. 53, pp. 2539 – 2547 2006

[48] Tao Yang; Mingkun Gao; Suntao Wu; Donghui Guo,”A new reuse method of analog circuit design for CMOS technology migration.” Anti-Counterfeiting Security and Identification in Communication (ASID), pp.112-115.2010

[49] Daniel P.Foty, ”MOSFET MODELING with SPICE: principles and practice.” 1997 by Prentice Hall PTR Prentice-Hall, Inc.

[50] A. I. A. Cunha, M. C. Schneider, and C. Galup-Montoro, ”An MOS Transistor model for analog circuit design,” IEEE J. Solid-State Circuits, vol. 33,pp.1510-1519, October1998.

[51] W. M. Sansen, F. O. Eynde, and M. Steyaert, “A CMOS temperature compensated current reference,” IEEE J. Solid-State Circuits, vol. SC-23, pp. 821-824, June 1988.

Chapter -4 Low-power and Low-speed SAR ADC Design Methodology with LDE-aware LCR

This section presents that LDE-aware LCR is applied and verified by a low-power and low-speed successive-approximation register (SAR) analog-to-digital converter (ADC). This SAR ADC is popularly used in medical implant systems. In the application of low-power and low sampling rate ADCs, reduction of leakage current mainly contributes to the total energy consumption. Focusing on the shallow trench isolation (STI) as a layout dependent effect, we propose LDE-aware LCR to reduce a significant junction leakage current through the trench isolation. We design a 11-bit SAR ADC with our technology in 65nm CMOS process, which employs a dual supply mode (0.8V for analog and 0.35V for digital). From results, our SAR ADC achieves DNL of +0.73/-0.62LSB, INL of +0.37/-0.79LSB, SNDR of 55.93 dB, SFDR of 64.47dB and ENOBs of 9-bits at 5KS/s sampling rate. The total power consumption is estimated as 38.6-nW. Comparing with the same design without LDE-aware leakage reduction, we attain the 30% total energy reduction.

4.1 Design Trend of Analog-to-Digital Converters and Design Gist

Fig.4.1 shows the resolution and sampling frequency for all ADCs published in key technical conferences in this field (ISSCC and VLSI) between 1997 and 2012 [52]. The plot shows the trend that increasing sampling frequency goes with decreasing resolution. Of the classical architectures, $\sum\Delta$

converters dominate the high resolution and low sampling frequency region, flash and folding ADCs have the highest sampling frequency but with the lowest resolution, successive-approximation-register (SAR) converters are used for low-to-medium speed and medium-to-high resolution applications, and pipelined converters are used for applications that require medium-to-high speed and resolution.

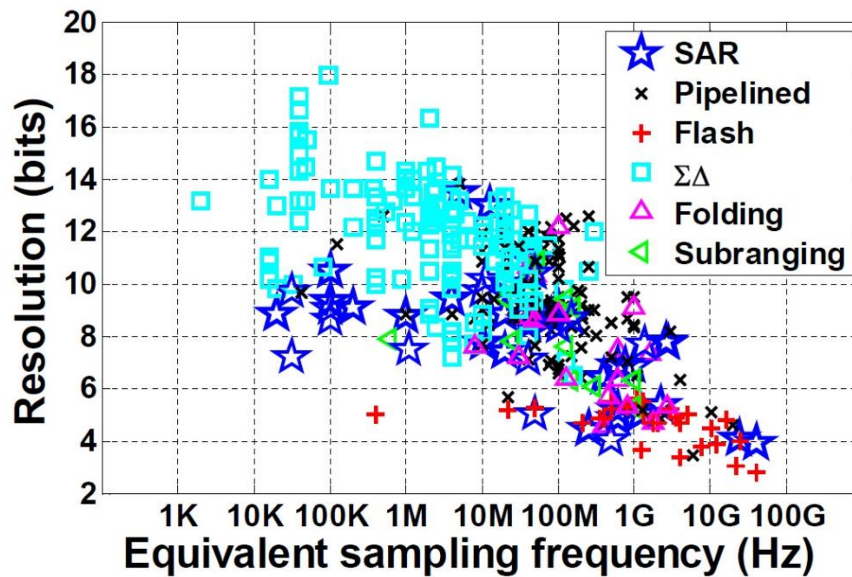


Fig.4.1 A plot of the resolution versus the input sampling frequency for recent published analog-to-digital converters in ISSCC and VLSI (data adopted from [52]).

Capacitor array successive-approximation-register (SAR) ADCs were introduced in 1975 by McCreary et al. [53] and have been extensively used for medium-speed applications. A conventional SAR ADC includes a digital-to-analog converter (DAC) driving a comparator. The comparator output is then processed by the digital control logic, which in turn feeds back a control signal to the DAC. This feedback logic is performing a binary search to find the correct digital output bits to minimize the difference between the voltage on the DAC and the analog input. The DAC is typically

composed of binary-weighted capacitors, which also serve as the input sampling capacitor. A sub-DAC can be used to avoid large capacitor values and enable high resolution implementations. The architecture has very high energy efficiency since other than the comparator, the remaining blocks only consume dynamic power. One drawback of the SAR architecture is that it takes multiple clock cycles, usually the same as the number of bits, to generate an output. This has made it difficult for the SAR architecture to run at sampling rate more than 5MHz in the past. Digital scaling helps improve the speed of CMOS technologies, now making SAR a viable option for higher speed applications. Moreover, scaling issues affecting other architectures are not present in SAR ADCs because of its high digital composition.

A widely adopted Figures of Merit (FOM), also called Walden's Figures of Merit [54], that incorporates resolution, speed and power consumption in order to provide a platform for energy efficiency comparison is shown below:

$$\boldsymbol{FOM_1 = \frac{P}{2f_{sig} \times 2^{ENOB}}}, \qquad \textbf{(Eq.4.1)}$$

$$\boldsymbol{ENOB = \frac{SNDR - 1.76}{6.02}}. \qquad \textbf{(Eq.4.2)}$$

Where P is the total power consumption, ENOB is the effective number of bits, defined in Equation 1.4, and $\boldsymbol{f_{sig}}$ is the input frequency of the signal. SNDR is the signal-to-noise-distortion ratio in dB measured with a sinusoidal input. This FOM is intended to provide a measure of how much energy is required to perform a conversion step, expressed in pico-joules (pJ) per conversion step. The development of this FOM is mostly based on empirical data after surveying a large number of ADCs in academic

publications or commercial ADCs. This metric is created under the assumption that power tends to scale linearly with the input frequency and SNDR. It allows designers to compare energy efficiency across ADCs operating under different conditions. However, this metric suffers from an important limitation. In a higher accuracy ADC that is 10 bits or more, the resolution is mostly limited by thermal noise, which is in the form of $\sqrt{kT/C}$ In order to increase the resolution by 1 bit (or SNR by 6dB), C has to quadruple. If the operating frequency is kept constant, the power consumption has to be increased by a factor of four for an improvement of a factor of two in resolution. This implies that improving the resolution by 1 bit automatically worsens the FOM by a factor of two.

To address these limitation due to thermal noise, a modified FOM is proposed by [55] as shown in Eq.5.3:

$$\boldsymbol{FOM_2 = \frac{P}{2f_{sig} \times SNTR^2}}. \qquad \textbf{(Eq.4.3)}$$

Where SNTR is the signal-to-thermal-noise ratio. In the absence of distortion and quantization noise, SNTR = $\mathbf{2^{ENOB}}$. Since the sampled thermal noise of an ADC is in the form of $\sqrt{kT/C}$, the square of signal-to-thermal-noise-ratio

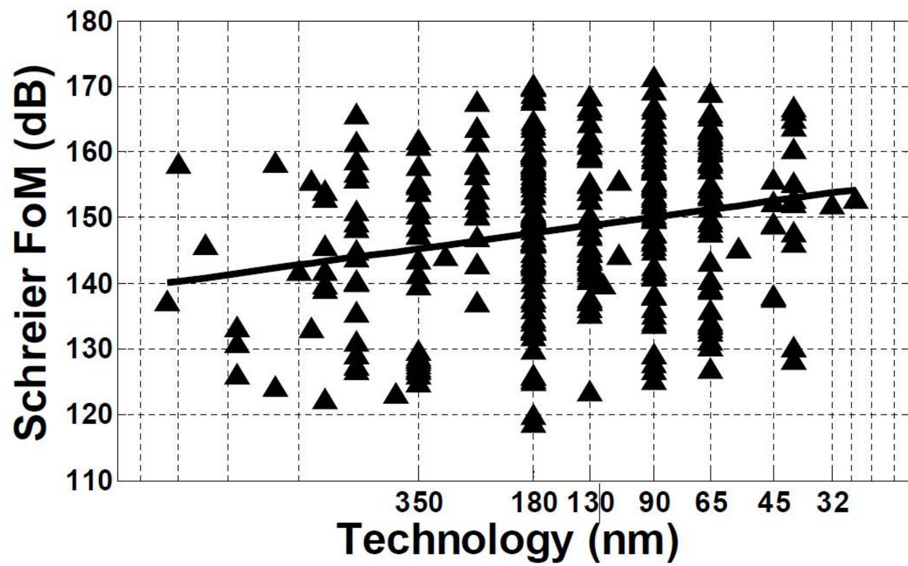


Fig.4.2 The Schreier Figure of Merit ($\boldsymbol{FOM_3}$) versus CMOS process nodes from 1um to 28nm of state-of-the-art ADCs published at ISSCC and VLSI Symposium (data adopted from [52]). Even though technology scaling does not directly benefit analog integrated circuits, steady improvement in conversion energy efficiency, $\boldsymbol{FOM_3}$, is still shown. This trend is the result of using more digital friendly architectures in recent designs.

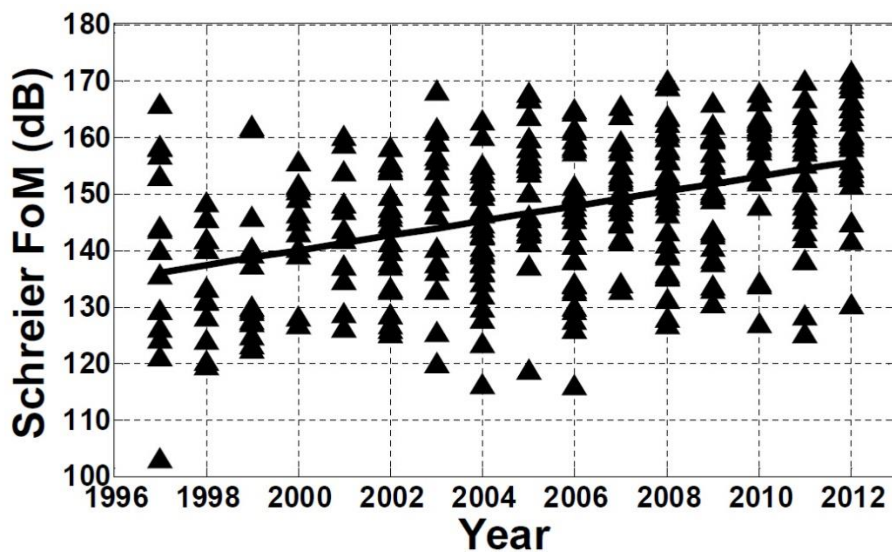


Fig.4.3 The conversion energy efficiency FOM_3 from year 1997-2012 (data adopted from [52]). This trend emphasizes the importance of energy efficiency in recent designs.

($SNTR^2$) is thus proportional to C. In other words, at a fixed sampling frequency, the increase in power requirement is the same as the increase in SNTR2, making the overall FOM_2 constant. This makes FoM2 more suitable for comparing high-accuracy ADCs that are thermal-noise limited.

Another variant of a figure of merit is named the Schreier FoM, listed below as FOM_3 [56]. It inverts the previous FOM_2 and expresses the term in dB, increasing in values for higher performance ADCs at the same frequency.

$$FOM_3 = SNDR_{dB} + 10log(\frac{2f_{sig}}{P}). \qquad \textbf{(Eq.4.4)}$$

Fig4.2 and Fig.4.3 plot Schreier's FOM (FOM_3) versus technology and years for the state-of-the-art ADCs published at ISSCC and VLSI Symposium between 1997 and 2012 [52]. It shows a general increasing trend; on average, FOM_3 increases by 1.3dB per year. This improvement in FOM_3 could partially be attributed to the use and invention of more digital friendly ADC architectures.

Fig4.4 and Fig.4.5 show the values of plotted against sampling frequency and resolution, respectively. The best converters can achieve figure of merit at tens of femto-joule per conversion step; however, these ADCs tend to be at resolution lower than 10 bits and sampling rate less than a few mega-samples per second. In terms of energy efficiency, Fig.4.4 shows that the SAR architecture dominates the figure of merit over all other

architectures for all sampling frequencies between 10KS/s and 1GS/s. When sampling frequency increases, it becomes more difficult to achieve the same energy efficiency as for designs with lower frequencies. These so-called "high-speed" ADCs rely more heavily on the speed capability of the underlying transistors. To run at faster speed, extra power needs to be burnt to hit the performance targets.

In terms of resolution, Fig.5.5 shows another interesting trend. Converters with ENOBs between 6 and 10 bits are able to achieve the best Walden's FOM ($\boldsymbol{FOM_1}$). This window is the "sweet spot" for achieving energy efficient designs, and we refer to this as the Energy Efficiency Window (EEW). The EEW is especially suited for designs that are used for the battery-sensitive portable devices. For resolution lower than 6 bits, the

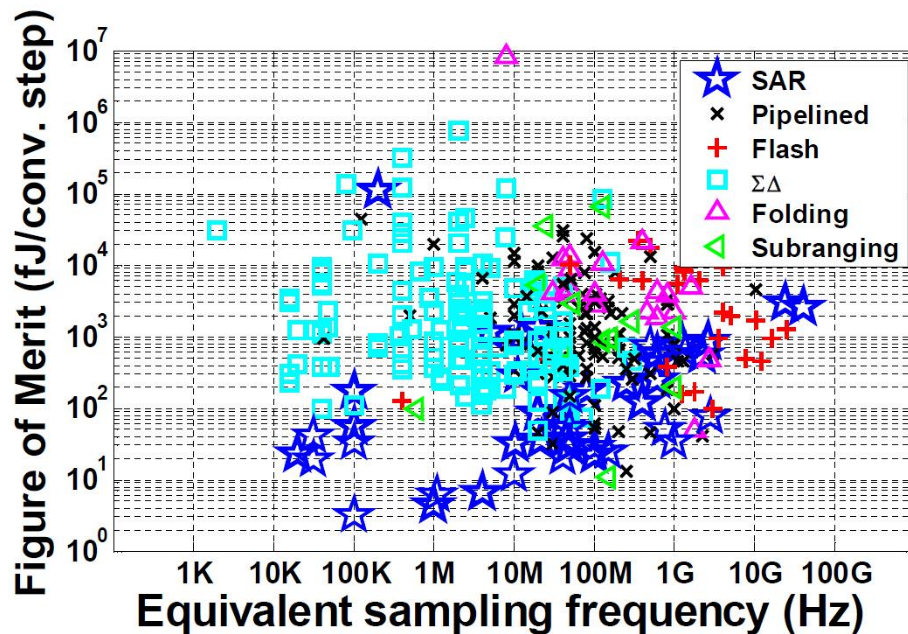


Fig.4.4: Walden's FOM versus sampling frequency of state-of-the-art ADCs published at ISSCC and VLSI Symposium (data adopted from [52]).

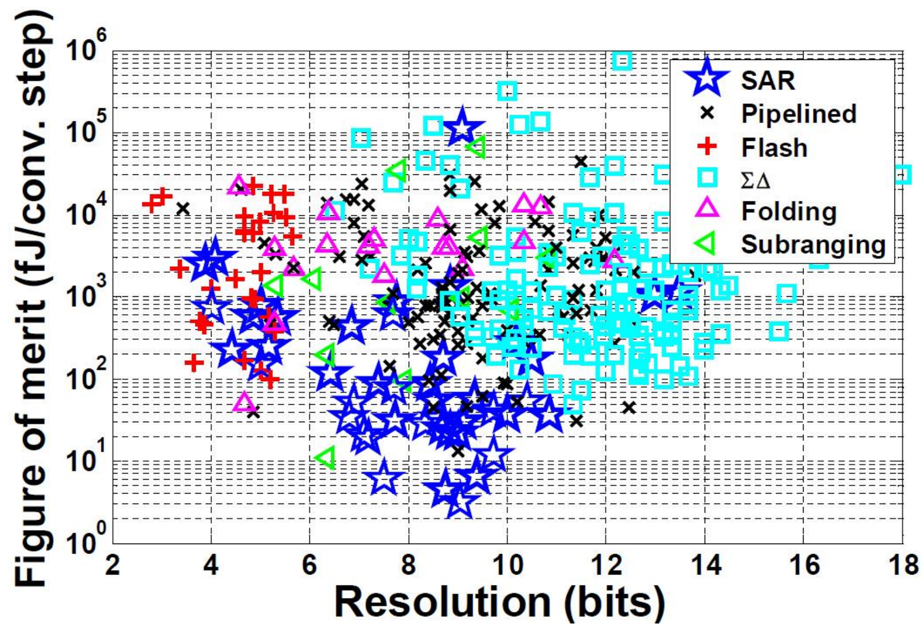


Fig.4.5: Walden's FOM versus resolutions of state-of-the-art ADCs published at ISSCC and VLSI Symposium (data adopted from [52]).

design is typically targeted for very high-speed ADCs, where, as explained before, it is difficult to improve energy efficiency due to technology limitations. For resolution more than 10 ENOBs, thermal noise degrades $\boldsymbol{FOM_1}$. Because the designs are noise-limited, various oversampling techniques are needed to lower the effective thermal noise within band. These techniques are typically hard to implement in an energy-efficient fashion due to its high oversampling ratio. As shown in Fig.4.5, within the EEW, the SAR architecture again dominates the energy efficiency over other architectures. One of the goals in this thesis is to expand the width of the EEW to allow energy efficient design in higher resolutions that have more than 10 ENOBs.

Because the goal of this thesis is to advance medium resolution (11bits) low frequency (5KS/s) ADC design, looking at the energy efficiency using FOM will give us additional perspective. To meet this competitive energy

efficiency at ENOB more than 8 bits, we choose to explore the SAR architectures because of its high energy efficiency, small feature size and good digital compatibility. Being free of precision analog circuitry (besides the comparator), SAR ADCs scale very well with technology as they are less affected by the degraded intrinsic gain and shrunk voltage headroom than OPAMP based architectures such as pipeline ADCs. They can take better advantage of the speed and energy efficiency in deeply scaled CMOS processes.

4.2 SAR ADC Structures and Design Consideration

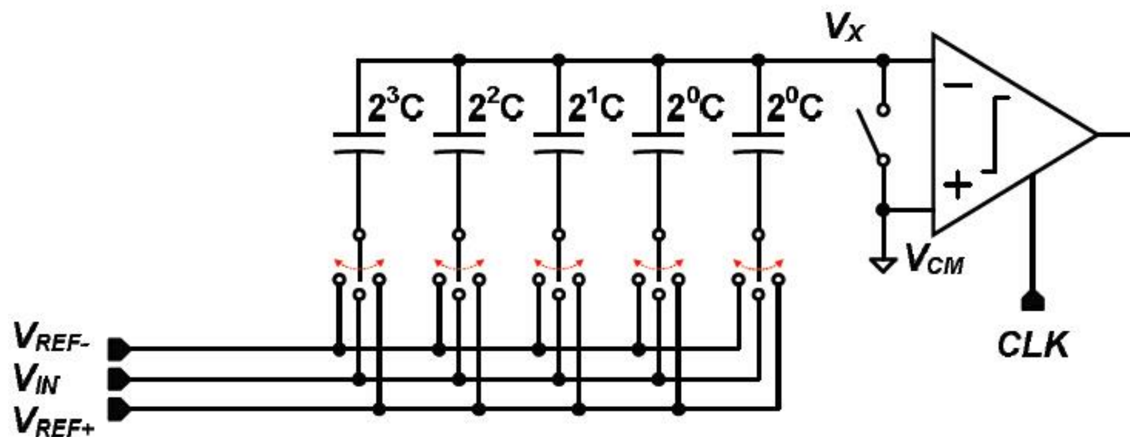


Fig.4.6 Structure of a binary weighted capacitor (BWC) DAC

DACs have many names according to the criteria for its classification: radix, grouping method of DAC elements, kinds of DAC elements and hybrid. If the criterion is the radix. DACs are classified as binary or non-binary, if it is the grouping method of DAC elements. DACs are done by binary controlled DAC or thermometer controlled DAC, if it is the kinds of DAC elements. DACs are categorized by CDAC or RDAC, and finally if it is hybrid, DACs are grouped as "R and R'-based-DAC, 'C and R'-based-DAC, or 'C and C'

based-DAC. The most commonly used four DAC structures are shown from Fig.4.6 to Fig4.9 about 4b resolution.

Fig.4.6 shows the most basic BWC DAC structure. During A/D conversion, the comparator minus input of $\boldsymbol{V_X}$ is calculated by the charge conservation law, because charges at the output node of the DAC are frozen. The BWC DAC has next features: less number of switches and logics, but it has relatively large DNL peak. The total capacitance is defined as 2^NC.

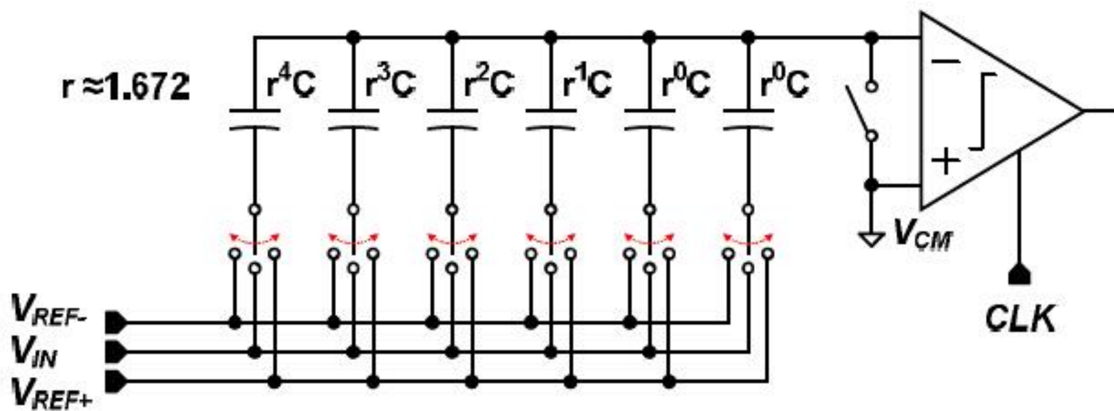


Fig.4.7 Structure of a sub-radix non-binary weighted capacitor (NWC) DAC

Fig.4.7 shows the example of a NWC DAC structure. The comparator minus input of $\boldsymbol{V_X}$ is calculated by the same manner as mentioned above. This DAC must have more than 1 or 2 capacitor elements, too. As the sub-radix DAC requires more than 1 or 2 phases to get a proper digital output. This is the concept of the redundancy, and if an ADC obtains 4b ENOB through 5 times A/D conversions, the radix (r) is given by following Eq. (4.5):

$$\boldsymbol{\frac{r^5-1}{r-1} = 2^4 - 1} \qquad \textbf{(Eq.4.5)}$$

This equation comes from the fact that the total capacitance of a BWC DAC is the same with that of a NWC DAC. Frankly speaking, the NWC DAC was not practical due to the non-integer capacitor ratio; however, some designers have tried to overcome the capacitor mismatch problem by applying the LMS algorithm [56]-[58]. This NWC DAC results in a reduction in DAC settling time, even though sacrificing conversion time due to the additional extra comparison cycle(s).

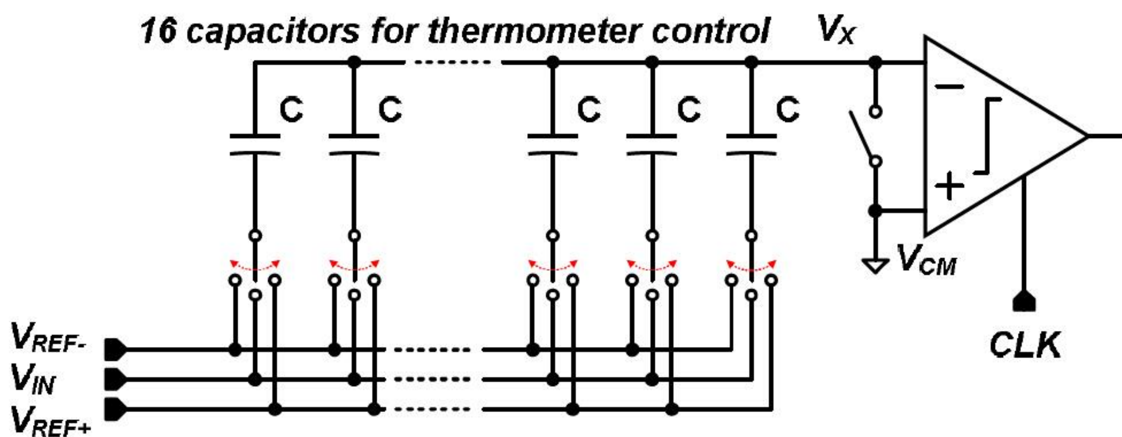


Fig.4.8 Structure of a thermometer controlled NWC DAC

In fact. when the Kuttner proposed the non-binary SAR ADC with a capacitor DAC firstly at 2000's ISSCC [59], the DAC structure he implemented is the following Fig.4.8. All elements of the DACs are controlled one by one, and its sub-radix is defined by digital ROM. This DAC structure has the same INL plot as the BWC DAC.

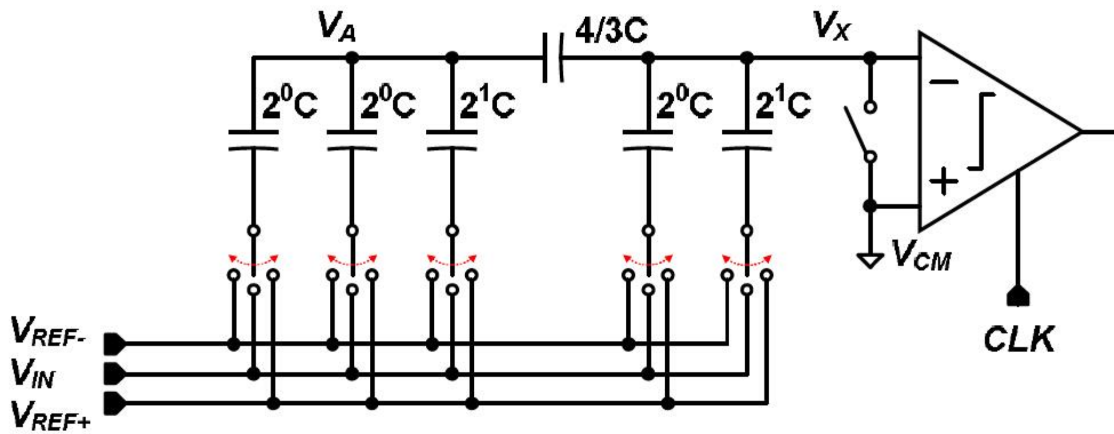


Fig.4.9 Structure of a Bridge capacitor (BC) DAC

The BC DAC as shown in Fig. 2-4 is one type of the hybrid DAC mixing two capacitor DACs. There are two elements at the MSB side and also two elements at the LSB side. Therefore it is called 2b+2b hybrid structure. This structure was also not feasible for more than l0b ADC since the mismatch of the weighting factors between the smallest capacitor of MSB side and all capacitors at the LSB side is large which comes from fractional bridge capacitor and which is induced by a parasitic capacitor at the VA node. However, the growing interest in high resolution and high speed SAR ADCs makes the BC DAC be considered again because this DAC structure has small sample capacitance equivalently. With the calibration algorithms, the BC DAC has frequently been selecting to a 12b SAR ADC.

The capacitor DAC in a SAR ADC is usually used as sampling capacitor of S/H circuits. Therefore, total sampling capacitor must be designed to satisfy the KT/C noise requirement. The relation of signal power to KT/C noise is expressed by following Eq. (4.6) about the single ended sampling network when V_{PP} indicates peak-to-peak signal voltage

$$Signal - to - \frac{KT}{C} noise_{single-end} = 10log\frac{P_{Signal}}{P_{KT/C}} = 10log\frac{V_{PP}^2}{KT/C}$$

(Eq.4.6)

If the input sampling network is designed differentially, because signal amplitude is two times larger and noise power increases as two times, the relation is modified as follows:

$$\boldsymbol{Signal - to - \frac{KT}{C} noise_{diff} = 10log\frac{(2V_{PP})^2/2}{2 \cdot KT/C} = 3dB + 10log\frac{V_{PP}^2/2}{KT/C}}$$

(Eq.4.7)

Using this equation, we can obtain minimum capacitance to satisfy the KT/C noise requirement when the resolution and signal power are given.

To know the size of a unit capacitor, the designer must find the standard deviation value (one sigma = σ) which meets certain yield (ex. 99.9%) at first. The specific sigma value has the designers choose the unit capacitor size with the mismatch information provided by the foundries as shown in Fig.4.10. The sigma value can be obtained by a behavioral MATLAB simulation [60].

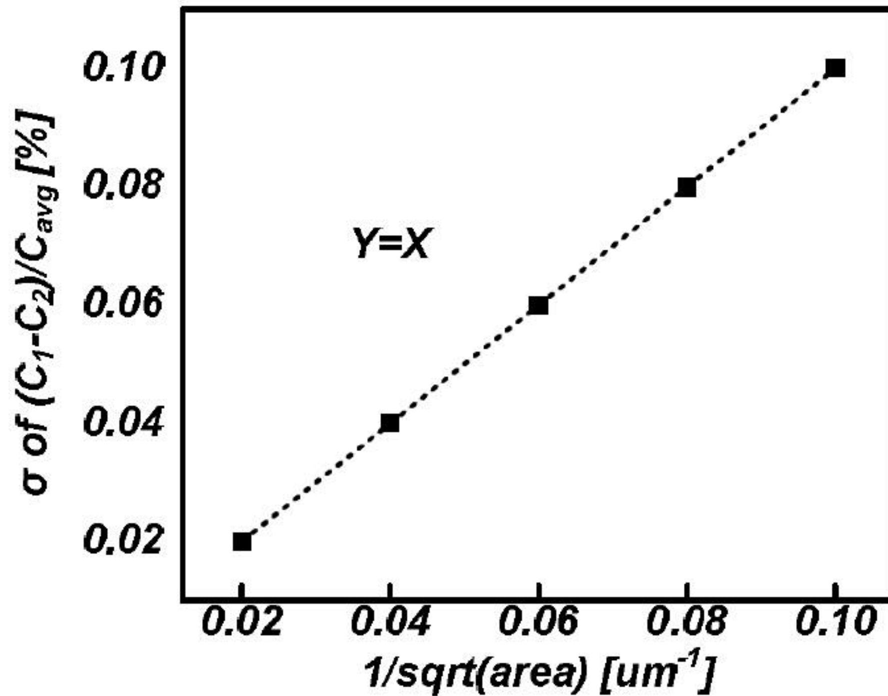


Fig.4.10 Capacitor mismatch

4.3 SAR ADC Architecture Design with Considering LDE-aware LCR

4.3.1 Description of Proposed Overall Design Flows

For the practicality of LDE-aware LCR, we tackle the dedicated leakage reduction technique for the performance of a specific 11-bit SAR ADC. In the practical example of SAR ADC, we discuss the design procedure step by step and implement LDE-aware LCR in fundamental blocks. As shown in Fig.4.11, there are usually three steps to design the whole SAR ADC schematic. Since we focus on the low-speed SAR ADC, a loose per-conversion time requirement allows us to employ LDE-aware LCR in the third step, which are not suitable for high-speed ADCs. Our design flow consists of three steps:

- Step 1) is to verify specifications to clarify a proper ADC style.
- Step 2) is to decide an architecture from global view, especially focusing on a switching algorithm and a matching capacitor array.
- Step 3) is to implement LDE-aware LCR and low-power circuit in fundamental blocks.

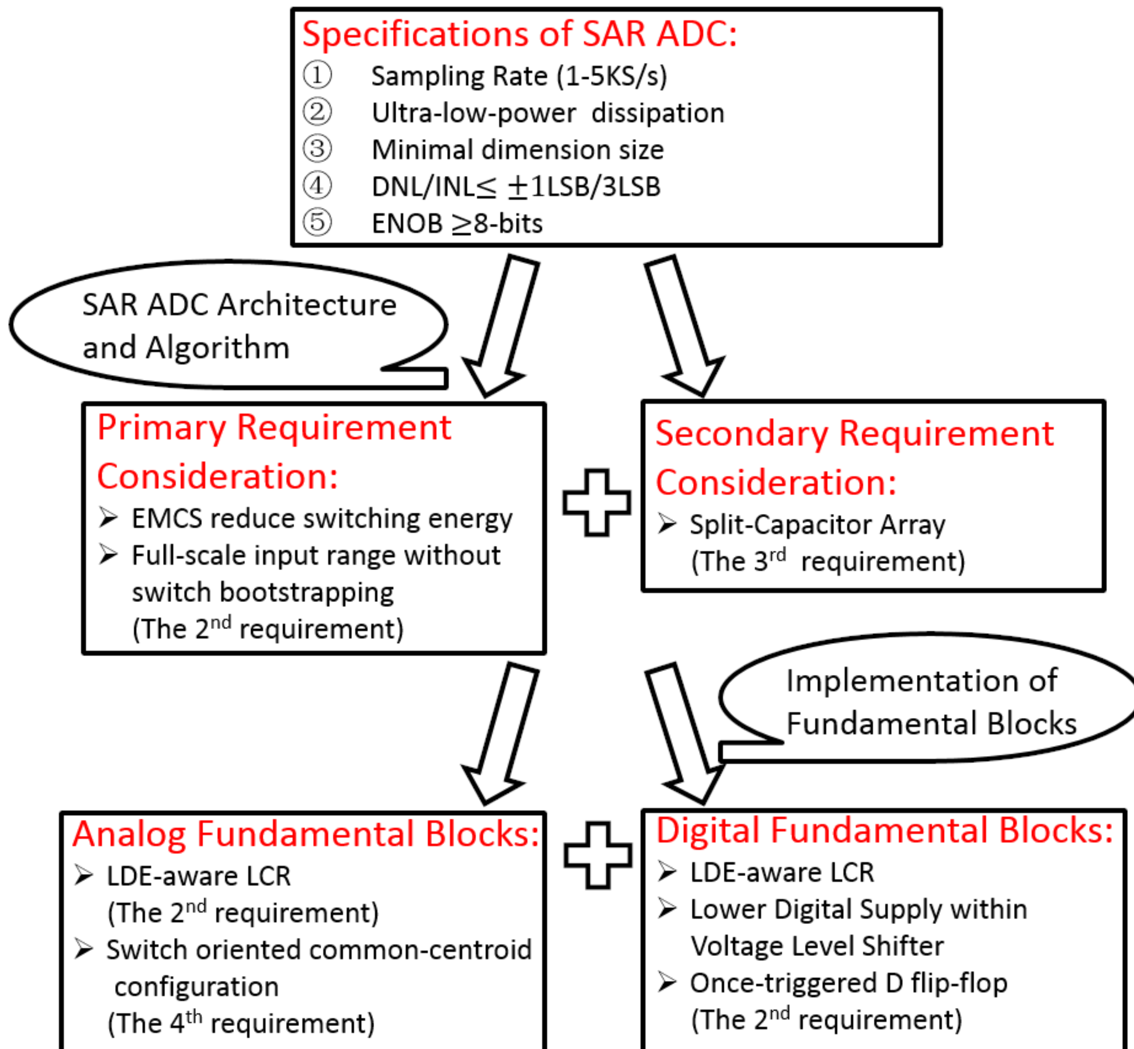


Fig.4.11 The Flow for Proposed Synthesized Low Power Design Strategy

In details, firstly, we check a matching between a given specification and an assumed SAR ADC. We cannot always synthesize circuits for any specification in the SAR ADC. We use a limitation of the specification in the

subsequent synthesizing steps. We notice that a typical SAR ADC has an advantage of the power efficiency and simple analog sub-circuits. The only two analog components are comparator and capacitive DAC. No static power is consumed if dynamic comparator can be used. Therefore, just with a simple architecture, we can meet the design requirements.

Next, we choose architectures of SAR ADC components according to the requirements. In this work, a first primary requirement is to reduce the power at ultra low-power level, and a second primary requirement is minimizing the dimension size as small as for the prospective feasible application of array sensor. That is, both requirements are objective to determine the architectures composing the SAR ADC. Referring the other requirements are regarded as constraints. The accepted DNL/INL are smaller than ±1LSB/3LSB, respectively. ENOBs should exceed 8-bits.

Since a switched capacitor equivalent resistance becomes larger at the low sampling rate, a switching sequence gives a large impact with respect to the power consumption in nano-watts level. An excellent switching procedure is essential for the power efficiency. Compared to the conventional switching sequences, the energy saving [61], monotonic [62], *Vcm*-based [63], early reset merged capacitor switching (EMCS) [64] and, partial floating [65] switching sequences can reduce switching energy by 69%, 81%, 90%, 91.4% and 94%, respectively . The EMCS provides an advantage in the best power efficiency and the simplicity of circuit operations. Adopting the top-plate sampling with a direct MSB decision procedure, we need not an extra switch bootstrapping anymore. Therefore, we adopted EMCS in this work.

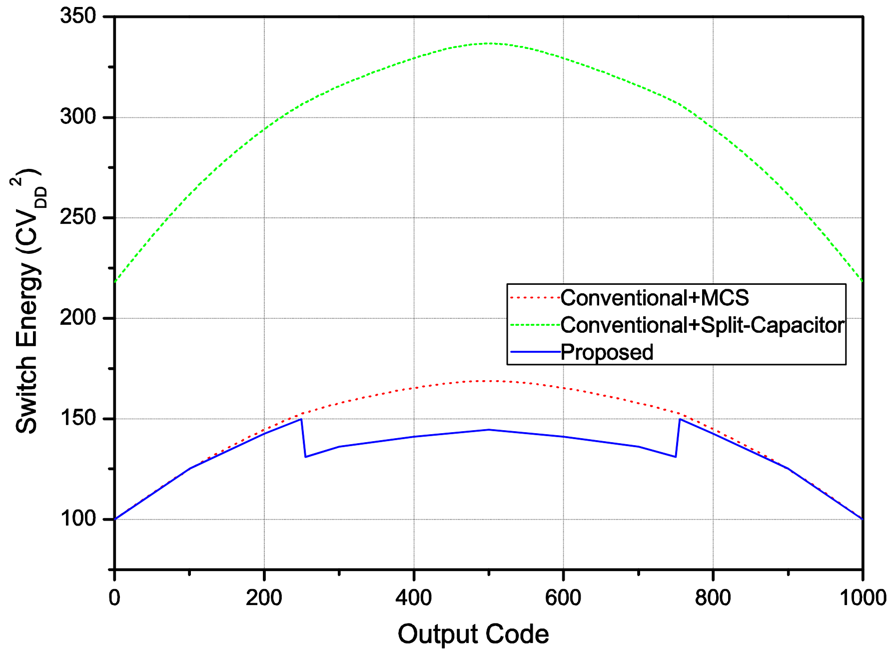


Fig.4.13 Switch energy versus output code for three different techniques

The switching procedure of the EMCs algorithm also serves a switching energy saving compared with the conventional one. Fig.4.12 demonstrates 3-bit examples of the conventional and proposed switching methods. Fig.4.12 (a) shows all possible conversions. The quantitative energy consumption of each switching phase is also shown in the figure. The conventional switching sequence is efficient when all the attempts are successful as in the upper case. However, the switching sequence consumes a lot of energy when attempts are unsuccessful as in the lower case. As seen in Fig.4.12 (b), the subsequent switching sequence of the proposed method is more efficient than that of the conventional one. For the second bit estimation, as indicated in Fig.4.12 (b) in the situation of MSB=1 (the upper case), the outputs of DACs finally settle to the following values:

$$\boldsymbol{C_{-}[0] = C_{-}[1] = V_{DD} - V_{in+},}$$

$$C_+[0] = C_+[1] = V_{DD} - V_{in-},$$

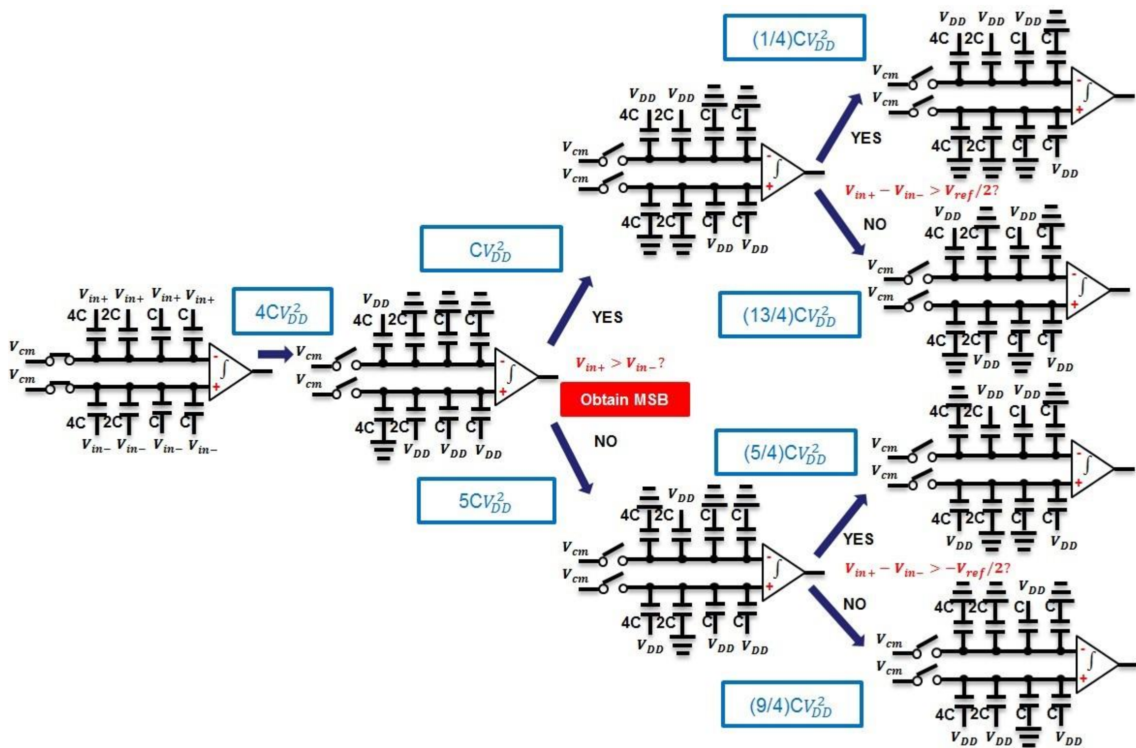
$$C_-[2] = V_{DD} - V_{in+} - (V_{DD} - V_{CM})\frac{2C}{4C},$$

$$C_+[2] = V_{DD} - V_{in-} - (V_{GND} - V_{CM})\frac{2C}{4C}. \quad \textbf{(Eq.4.8)}$$

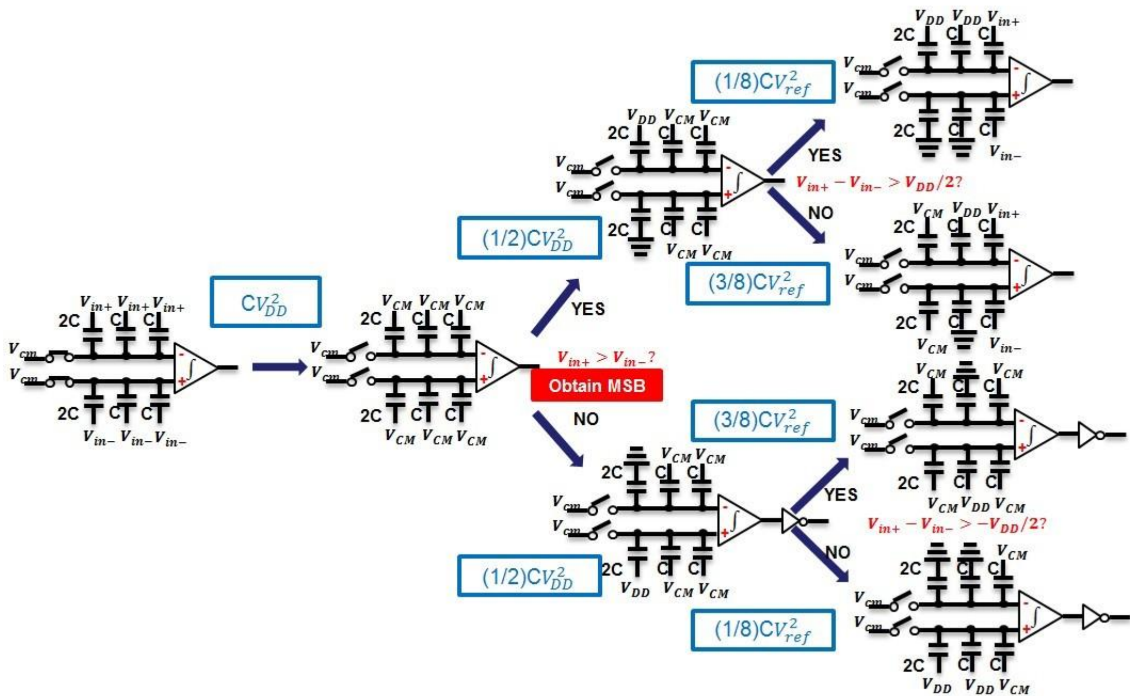
The necessary energy for the pre-charge transition is

$$E_{charge} = \frac{1}{2}V_{CM} \times 2C[(C_-[2] - C_-[1]) + (C_+[2] - C_+[1])] = -\frac{1}{4}CV_{DD}^2. \quad \textbf{(Eq.4.9)}$$

Comparing the power dissipation of these methods, it is obvious that, for every possible code, the EMCS SAR needs equal or lower energy. For a 3 bit case, the conventional switching procedure consumes $\mathbf{8.75}CV_{DD}^2$, while the proposed method does only $\mathbf{1.75}CV_{DD}^2$. Thus, the proposed method requires 80% less energy than that of the conventional one. Fig.4.13 compares the proposed method with two other conventional methods with double input range combining with different techniques for switching energy saving, called the Split-Capacitor method [6] and the merged capacitor switching (MCS) method [4]. In the figures, we assume that an equivalent unit capacitor is used for 9 bit capacitor arrays. The EMCS algorithm grants about 10% benefit than the MCS method overall output codes.



(a) Conventional switching procedure



(b) EMCS switching procedure

Fig.4.12 3-bit examples of the conventional and proposed switching methods

Taking the minimal dimension size into account, the split-capacitor array is a better choice for the smaller geometric size of capacitor array [65] [68] [69]. In the preliminary works, we notice that if the low-speed ADC utilize proper unitary capacitance and matched capacitor array layout, the ADC can be synthesized to 11-bit without digital or analog calibration. Herein this design strategy, we adopt a combination of EMCS and split-capacitor array as our SAR ADC architecture.

In the third step, we give the further implementation of the low-power design in the fundamental block level. In these blocks, all transistors in the switches and comparator work in the moderate or weak inversion region. Hence the leakage current becomes to be a large issue. Since CMOS process has been scaled-down to 90 nanometers and below, instead of sub-threshold channel leakage, the junction leakage has been to be a major component for the total leakage. In details, when the shallow trench isolation (STI) dislocations are located with the depletion region of pn junction, anomalous junction leakage current could flow. Thus we consider to use LDE-aware LCR in the fundamental blocks. Furthermore, for the energy saving in digital blocks, we employ a low digital supply of 0.35V with level shifters [70]. Besides, we also adopt a low-power consumption once-triggered D flip-flop [71]. Hence, the power consumption is further reduced.

4.3.2 Description of SAR ADC Architecture

Fig.4.14 shows the block diagram of the proposed SAR ADC. The fundamental blocks are the comparator, the capacitive DAC, level shifters and the logic control circuit. A differential input architecture is employed to have a good common mode noise rejection, and charge injection impact is

also reduced. A top-plate-sampled mode is adopted to cut the total capacitance in half. The smaller input capacitance also enhances the dynamic performance of the ADC.

Since this ADC is fully differential, the operations of the two sides of the circuit are complementary. For the simplicity, by using 4-bits example of Fig. 4.14, only the two steps of sample cycle and charge-redistribution cycle are described below. At the sample cycle, the top-plates of the capacitors are charged $\boldsymbol{V_{in+}}$, and the bottom-plates are reset to the common mode voltage $\boldsymbol{V_{CM}}$. The first comparison between $\boldsymbol{V_{in+}}$ and $\boldsymbol{V_{in-}}$ also made in this mode,

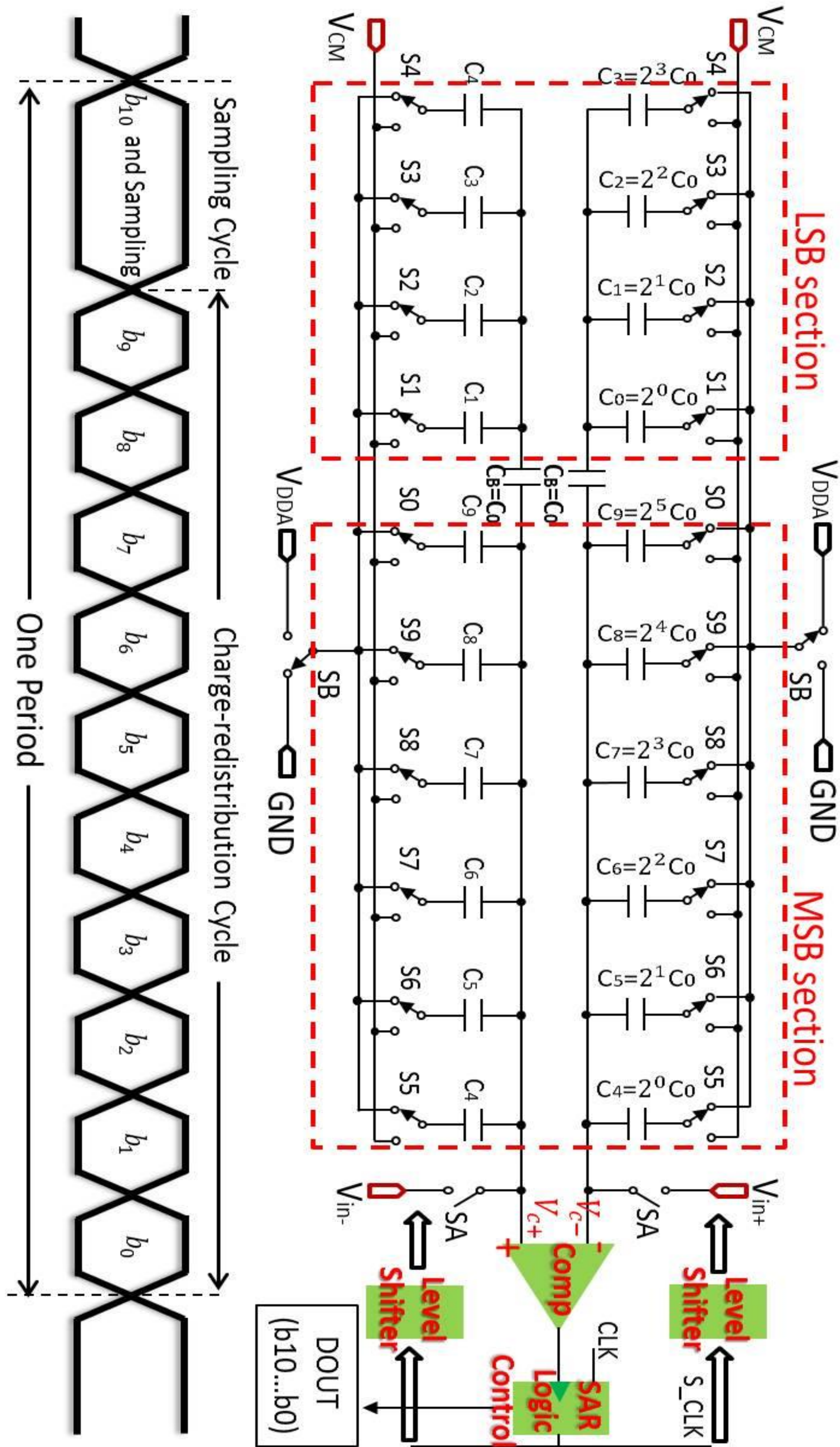


Fig.4.14 The proposed SAR ADC architecture

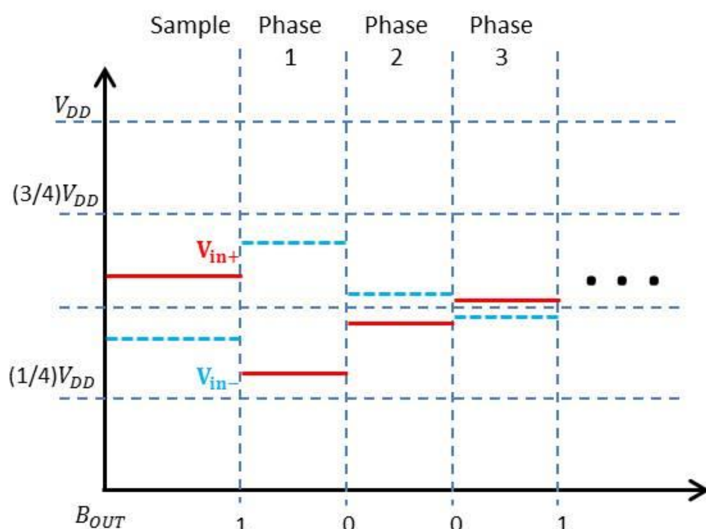


Fig.4.15 4-bits example waveform of the switching procedure

thus the polarity of the input is determined with the single comparator. Next, at the first step of charge-redistribution cycle, switch SA is opened. The bottom-plates charge of DAC is determined by the result of the first comparison. In the example, $\boldsymbol{V_{in+}}$ is higher than $\boldsymbol{V_{in-}}$, the output of the most significant bit (MSB, i.e. b10) is 1, $\boldsymbol{V_{DDA}}$ and GND switch to $C9$ of each $\boldsymbol{V_{in+}}$ side and the $\boldsymbol{V_{in-}}$ side, respectively, subtracting totally $\boldsymbol{V_{DDA}/4}$ from the differential input. The EMCS algorithm behaves different discharge in the reaming steps. Herein, b9 is 0, $C9$ is charged in the same manner with the DAC weight decreasing by a factor of 2. But $C9$ is reset to $\boldsymbol{V_{CM}}$, not like the conventional method to exchange switching of $\boldsymbol{V_{DD}}$ and ground (GND) to correct a previous virtual ground overshoot. The ADC repeats this procedure until the least significant bit (LSB) is determined. Fig. 4.15 is a 4-bits example waveform in the comparator inputs $\boldsymbol{V_{C-}}$ and $\boldsymbol{V_{C+}}$. In the sampling cycle, input signal is directly sampled and compared. In this example, the first comparison result is 1 (high). In the next steps of charge-redistribution cycle, the potential of $\boldsymbol{V_{C-}}$ and $\boldsymbol{V_{C+}}$ is to execute above

description procedure. And comparator outputs 0,0,1 to drive the SAR logic control.

The weights for MSB and LSB sections are given by:

$$W_{i,M} = \frac{C_i}{C_{MSB}+C_B \| C_{LSB}} V_{ref}, i = 4, 5, \ldots, 9,$$

$$W_{i,L} = \frac{C_i}{C_{MSB}+C_B \| C_{MSB}} \cdot \frac{C_B}{C_B+C_{MSB}} V_{ref}, i = 0, 1, 2, 3. \quad \textbf{(Eq.4.10)}$$

Where C_{LSB} and C_{MSB} are the total capacitance in the LSB and MSB sections, respectively. The MSB section has more than 2-bits binary-weighted capacitor array compared to the LSB section, thus to achieve good capacitor matching effect. If LSB and MSB sections have identical binary-weighted capacitors, and C_b matches the unitary capacitor C_0, the weights enables perfect linearity to this ADC. The nonlinearity of the SAR ADC is caused by the nonlinear weights due to the capacitor mismatch and parasitic. In high-speed ADCs, capacitor redundancy is utilized to meet the linearity requirement without degrading the speed. However, the calibration procedure calls for additional power dissipation. Benefiting from the low-speed requirement, the nonlinearity factors caused dynamic performance degradation is much smaller than high-speed ADCs, and dynamic performance is accepted in the design specifications. Therefore, the energy is saved with eliminating additional calibration devices.

4.4 Implementation of LDE-aware LCR in switch design

A sampling switch has to be carefully designed with consideration of linearity performance, since non-linearity in the sampling network contributes to overall ADC linearity errors. There are two kinds of switches: top plate sampling switch and bottom plate sampling switch

4.4.1 Top plate sampling (TPS) switch

Fig.4.16 shows a NMOS top plate sampling. This is called a "top plate sampling switch, because this circuit samples the input at the moment when the switch on top plate of sampling capacitor is turned off. When $\boldsymbol{\Phi_{sam}}$ is high, the NMOS switch operates at $\mathbf{V_{GS} = V_{DD} - V_{IN}}$, and the signal transfer function (STF) is expressed by the following Eq. (4.11):

$$\frac{V_{OUT}}{V_{IN}} = \frac{1}{1+S\cdot R_{on}(C_{sam}+C_j)}. \qquad \textbf{Eq. 4.11}$$

If R_{on} is an ideal resistor, the above STF certainly is linear on input amplitude. However, since R_{on} is the function of input amplitude as expressed in Eq. (4.12, the signal ain is different according to its input level.

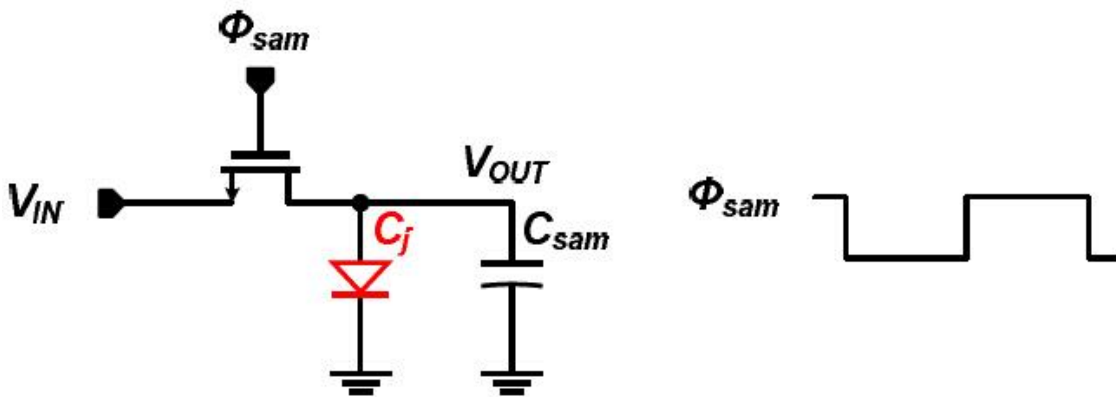


Fig.4.16 NMOS top plate sampling switch

This is the mam reason to injure its linearity on the TPS switch

$$\frac{1}{R_{on}} = \mu_n C_{OX}\frac{W}{L}(V_{DD} - V_{IN} - V_{th}). \qquad \textbf{Eq. 4.12}$$

There are the other four reasons; input dependent CLK feed through, input dependent channel charge injection, input dependent change of

threshold voltage ($\boldsymbol{V_{th}}$) by body effect, and input dependent capacitance of $\boldsymbol{C_j}$.

PMOS and CMOS top plate sampling switches can be identically understood with a NMOS TPS switch except for sampling the input of a higher common level.

4.4.2 Bottom plate sampling (BPS) switch

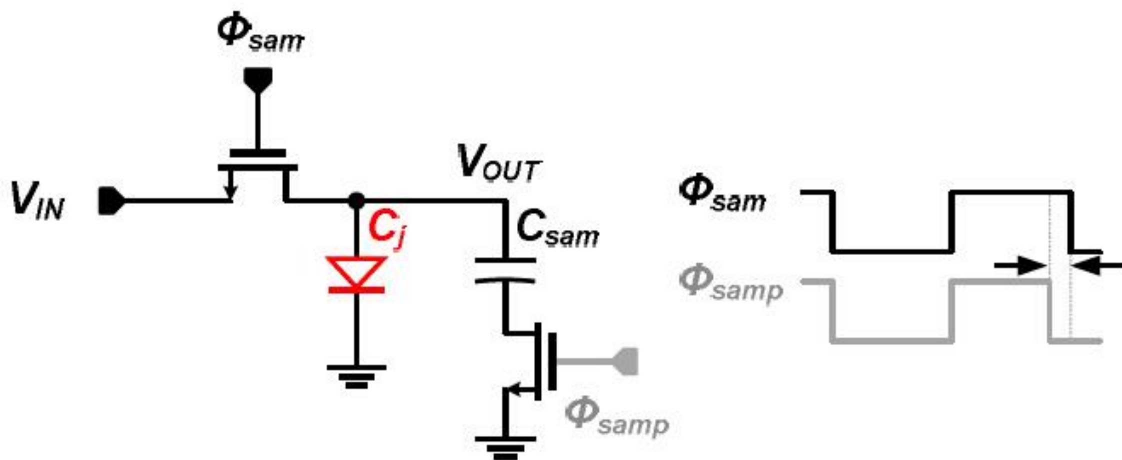


Fig.4.17 NMOS bottom plate sampling switch

Fig.4.17 shows a NMOS bottom plate sampling. This is called "bottom plate sampling" switch, because this circuit samples the input at the moment when the switch connected with bottom plate of sampling capacitor is turned off, not at the falling edge of $\boldsymbol{\Phi_{sam}}$.

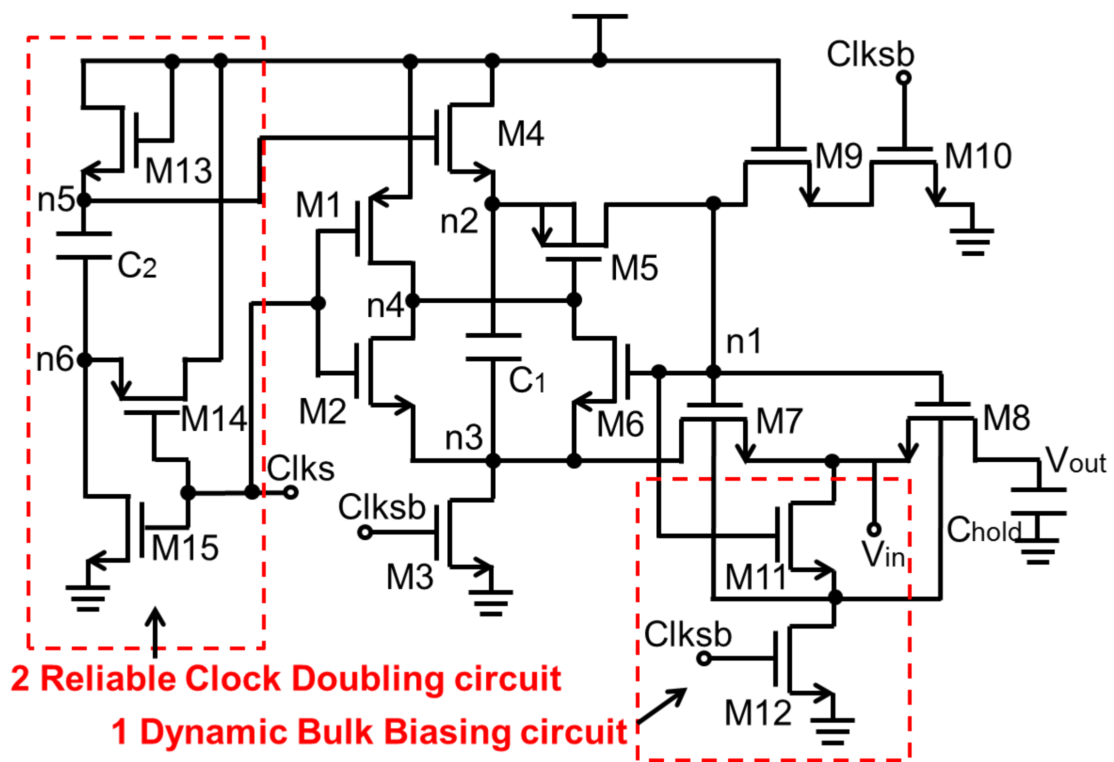
Many designers use the bottom plate sampling switch technique, because input dependent CLK feed through and channel charge injection problems are easily solved even through the prime CLK ($\boldsymbol{\Phi_{samp}}$) generator and one TR switch are additionally used. Note that the effects of nonlinear $\boldsymbol{R_{on}}$, varying $\boldsymbol{V_{th}}$, and nonlinear $\boldsymbol{C_j}$ on SFDR still remain as problems that must be solved.

4.4.3 A Linearity Bootstrapped Switch with Dynamic Bulk Biasing Design

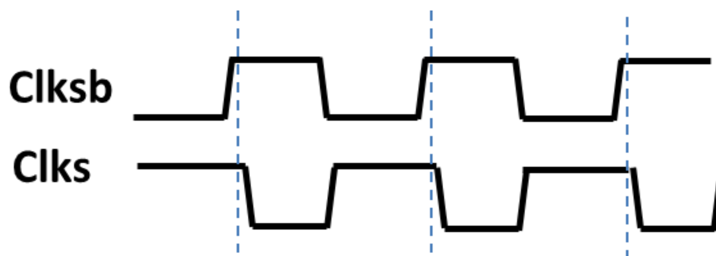
Fig.4.18 SAR ADC Scheme with Bootstrapped Switch

For achieving the design specification of high ENOBs, the front-end sample and hold circuit within high linearity performance should be a prevailing concern. Sampling switch non-linearity is mainly attributed to non-linear on resistance and associated parasitic capacitance which produce harmonic distortion when sampling high-frequency signals. The conventional bootstrapped switches in [72] has a merit of approximate constant on resistance because of the constant-overdrive-MOSFET approach. Since the bootstrapped switches with high linearity performance, both [63] and [73] adopted the bootstrapped switch as the sampling and holding circuit.

Although the bootstrapped switch can sampling the input signal



(a) The proposed bootstrapped switch circuit



(b)Clock Signal

Fig.4.19 The proposed bootstrapped switch circuit and clock signal

accurately, the non-linearity of internal input related transistors (except the booted transmission gate), and the charge rejection due to turning off some transistors slightly earlier in the overall circuit always cause an non-zero potential at the output when holding mode. As shown in Fig.4.18, the charge injection is corrected by introducing differential topology as it appears as a common-mode disturbance. Herein, a pair of cross-couple capacitors is used

to eliminate unequal charges in the comparison cycles. However, the charges introduced by two switches in a differential topology do not exactly cancel each other because of two reasons. The primary error is the load capacitance mismatch in the next stage of Digital-to-Analog converter (DAC) arrays. The secondary factor is the two input differential signals which are not equal to each other. Hence, for the purpose of obtaining much smaller nonzero holding biasing voltage, also achieving higher precision, an improvement for higher linearity of internal transistors is taken into consideration.

Fig.4.19 shows the proposed circuit and clock signal. The falling edge of CLKs has a time delay compare to the rising edge of $\boldsymbol{CLK_{sb}}$. In the initial of holding mode ($\boldsymbol{CLK_s}$=1.2V, $\boldsymbol{CLK_{sb}}$=1.2V), the bottom-plate of C1 is discharged through M3, and top-plate of C1 is discharged through M5, M9 and M10 simultaneously.

The whole circuit operation is discussed in a worst case of input signal equal to $\boldsymbol{V_{DD}}$. Transistors of M3, M4, M5, M7 and M10 correspond to five ideal switches. M8 is the main sampling switch whose gate is grounded through M9 and M10 during holding mode, hence, turning it off. During the same phase M13, M14, M15 and C2, forming a clock boosting circuit, drive transistor M4 which unidirectional charges capacitor C1. This solution provides clock doubling by ensuring the reliability constraint. None of the transistors in the proposed reliable clock doubling circuit have terminal voltages $\boldsymbol{V_{gs}}$, $\boldsymbol{V_{ds}}$ and $\boldsymbol{V_{gd}}$ exceeding the supply voltage of 1.2 V. Transistors M5 (connected to VDD through M1) and M7 isolate the sampling switch M8 when the capacitor C1 is being charged to VDD. The charged capacitor provides constant voltage of VDD between gate and source of M8 during sampling mode of the clock, thus also ensures a low on resistance independent of the input signal. During this mode, M2 pulls down

the gate of M5, turning it on and allowing the charge to flow from capacitor C1 to the gate of M6, M7 and M8. This turns all three transistors on. M7 also enables node n3 to follow the input voltage shifted by VDD, keeping the gate-source voltage constant regardless of the input signal. M9 reduces the $\boldsymbol{V_{ds}}$ and $\boldsymbol{V_{gd}}$ of M10 when $\boldsymbol{CLK_{sb}}$ is off and node n1 is at $\boldsymbol{2V_{DD}}$ for the worst case input of $\boldsymbol{V_{DD}}$. Transistor M6 ensures that the gate-source voltage across MP5 does not exceed $\boldsymbol{V_{DD}}$ by allowing the input voltage to appear at node n3 during sampling mode of the clock.

$$\boldsymbol{V_{n1} = V_{in} + \frac{C_1}{C_1+C_P} V_{DD}}. \qquad \textbf{Eq. 4.13}$$

The voltage at node n1 can be expressed in Eq. (4.13) with considering the parasitic capacitances attached to the top plate of C1. For the worst case input voltage of M7, which is an NMOS transistor, is required to conduct. The gate of this transistor is for this reason connected to the gate of M8 for bootstrapped voltage to ensure high conductivity by maintaining $\boldsymbol{V_{gs}}$ of the transistor equal to $\boldsymbol{V_{DD}}$ during sampling mode. In [73], it reports an option that to use PMOS instead of M4. But it would not have been possible to turn PMOS off during the sampling mode and it would have been a leaky switch. In this scheme M4 is a reliable switch. During holding mode, it charges capacitor C1 to $\boldsymbol{V_{DD}}$. During sampling mode, M4 has $\boldsymbol{V_{DD} - V_{thM4}}$ at the gate and $\boldsymbol{V_{DD} + V_{in}}$ at the drain. Even for the worst case, transistor M4 is off. Thus, the transistor reliability is ensured.

In the proposed dynamic bulk biasing circuit, both the bulk terminals of M7 and M8 is connected to the source of M11. When the sampling mode, M12 is turned off while M11 is turned on. The potential of drain and source of M11 is the same. Thus the substrate of M8 and M7 is equally connected

to the source of themselves. Even if $\boldsymbol{V_{in}}$ is higher than $\boldsymbol{V_{DD}}$, the PN junction between source and bulk of M8, M7 is not reversed. When the holding mode, M12 is turned on and M11 is turned off. This is helpful to discharge the parasitic capacitance of M7 and M8. In the reliable clock doubling circuit, during the sampling mode, capacitor C2 is pre-charged to a fixed value of $\boldsymbol{V_{DD} - V_{thM13}}$. During the holding mode, M14 connects the bottom-plate of C2. As a result the top-plate of C2 rises to $\boldsymbol{2V_{DD} - V_{thM13}}$. The gate potential of M4 toggles between $\boldsymbol{V_{DD} - V_{thM13}}$ and $\boldsymbol{V_{DD} - V_{thM13}}$.

Charge injection occurs due to unwanted charges injected into the circuit by turning off some transistors slightly earlier in the overall circuit. The output node $\boldsymbol{V_{out}}$ is the most sensitive from charge injection point of view. The charge injection is concluded from two aspects. The first dominating reason is that when M8 turns off, the channel charge, source-to-bulk parasitic capacitor charge and drain-to-bulk parasitic charge flow out to source and drain regions in a fast turning off time. The second reason is that the gate-drain overlap capacitance introduces variation to the output voltage during the holding mode. However, the charge injection from the overlap capacitance only has marginal effects on the output voltage. Herein, only the first dominating reason is considered.

It is supposed that the channel charge, source-to-bulk parasitic capacitor charge and drain-to-bulk parasitic charge of M8 flow equally in both directions toward source and drain. As in [25], the channel capacitance is calculate as

$$\boldsymbol{C_{channel} \cong 2/3 C_{OX} W_{M8} L_{M8}}. \qquad \textbf{Eq. 4.14}$$

The parasitic capacitance can be approximated as

$$C_{paraM8} \cong C_{diffusion} W_{M8} \times SA. \qquad \textbf{Eq. 4.15}$$

Where W_{M8} and L_{M8} is dimension of the transistor M8. C_{OX} and $C_{diffusion}$ are the gate oxide and diffusion unit area capacitance respectively. The drain and source of M8 is assumed has the same diffusion length of SA. The charge flowing to the output junction V_{out} for the bootstrapped voltage of $V_{dd} - V_{in}$ at gate terminal is given by:

$$\triangle Q = -(\frac{C_{channel}}{2} + C_{paraM8})(V_{dd} - V_{in}). \qquad \textbf{Eq. 4.16}$$

The change in voltage according to $Q = CV$ at the output node because of the charge is given as:

$$\triangle V = -\frac{(\frac{C_{channel}}{2} + C_{paraM8})(V_{dd} - V_{in})}{C_{hold}}. \qquad \textbf{Eq. 4.17}$$

Eq.4.17 proves that the dynamic bulk biasing circuit can decrease the charge injection effectively. Furthermore, in order to have a reduced charge injection at the output, either the size of the switch M8 should be small or the value of C_{hold} should be high. The bootstrapped switch has also an improved input voltage dependent signal distortion by keeping $V_{gs} = V_{dd}$.

The transistor size is listed in Table 1. The on-resistance of M8 is given by:

$$R_{on}(t) = \frac{1}{\mu_n C_{OX} \frac{W_{M8}}{L_{M8}} (V_{gsM8}(t) - V_{thM8}(t))}. \qquad \textbf{Eq. 4.18}$$

The sampling switch M8 is most critical from charge injection and low R_{on} point of view. If the width of the M8 is too large then it will result in an

Transistor	W[um]	L[um]	Transistor	W[um]	L[um]
M1	2	0.18	M9	2	0.3
M2	1	0.18	M10	2	0.18
M3	4	0.18	M11	1	0.18
M4	8	0.18	M12	1	0.18
M5	6	0.18	M13	2	0.18
M6	1	0.18	M14	4	0.18
M7	4	0.18	M15	2	0.18
M8	9	0.18			

Table4.1 TRANSISTOR SIZES FOR THE PROPOSED BOOTSTRAPPED SWITCH CIRCUIT

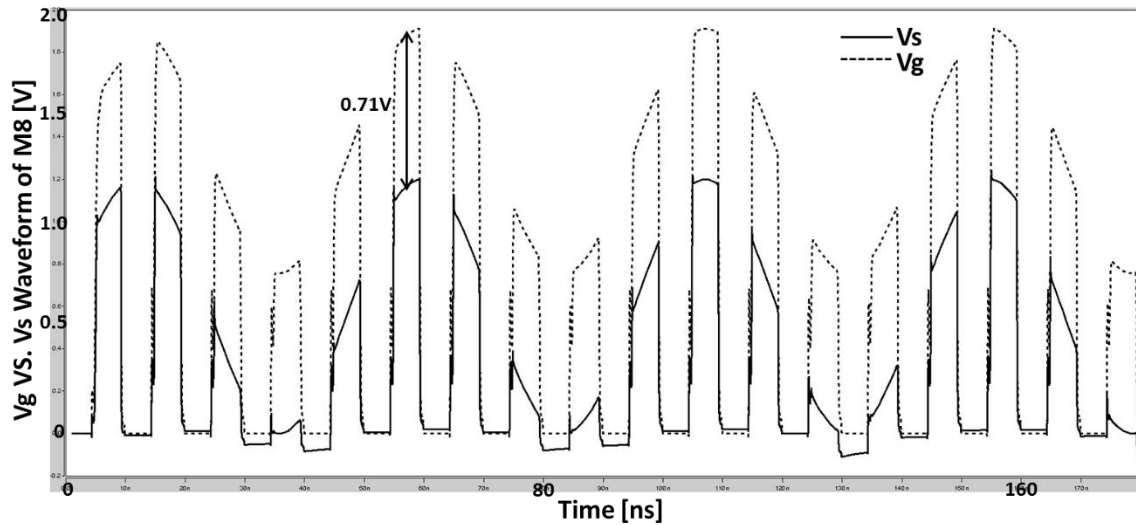
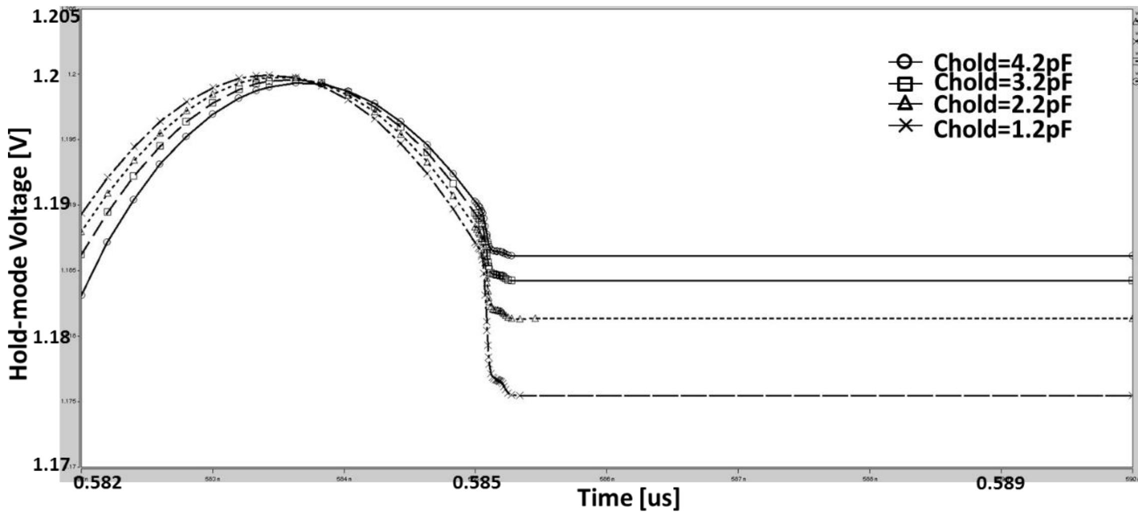


Fig.4.20 Simulation of Proposed Bootstrapped Switch Terminal Potential

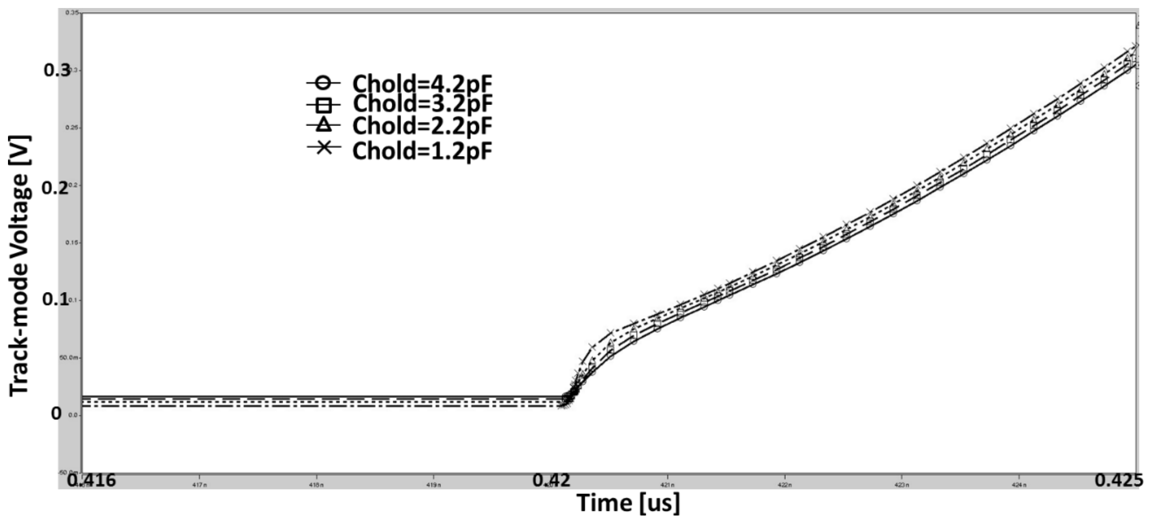
increased charge-injection at the output. On the other hand in order to have lower $\boldsymbol{R_{on}}$ value of the width of the sampling transistor should be larger. From simulations an optimum value of 9/0.18 um was selected for these trade-offs.

The transistor of M5 allows $\boldsymbol{V_{dd}+V_{in}}$ to appear as gate voltage for the sampling transistor. So the propagation delay of M5 should be as small

as possible. The same reason, for charging C1 fast enough, M4's width is four times of M13. All transistors in the design have a length of 0.18um (minimal length from design rule) except M9 whose length has been kept to 0.3um. It is helpful to improve the punch-through voltage of M9. The rest



(a) Charge injection during holding mode by sweeping different C_{hold} values



(b) Acquisition time during sampling mode by sweeping different C_{hold} values

Fig.4.21 Simulation of holding mode and sampling mode by sweeping different C_{hold} values

transistors in the bootstrapped switch are designed to have less propagation delay and thus avoiding unnecessary leakage paths because of different turn on times. C1 is 120fF and C2 is 80fF to have low KT/C noise. The capacitance of C_{hold} is determined by the whole SAR ADC design consideration. Followed by [70], a unit capacitor is selected by the thermal noise, capacitor mismatch and design rule. Herein, a unit capacitor is 17.76fF. For a 8bit calculation, the total loading capacitance of C_{hold} is 4.2pF.

The simulation of the bootstrapped switch is implemented in 65nm triple well process. Fig.4.20 shows the simulated gate and source voltage of M8. The discrepancy is preserving in a constant potential of 0.71V.

Fig.4.21 varies the value of C_{hold} from 1.2pF to 4.2pF. The effect of these variations on charge injection during holding mode and on acquisition time during sampling mode verifies the discussion before.

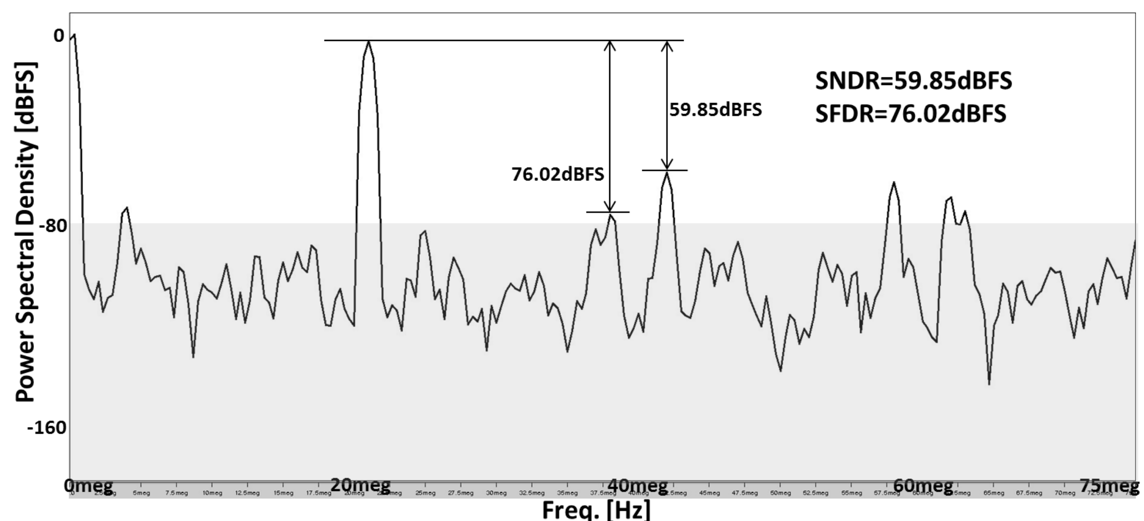


Fig.4.22 Simulated power spectrum of the proposed bootstrapped switch @ f_{in}=21MHz, f_{sample}=150MHz

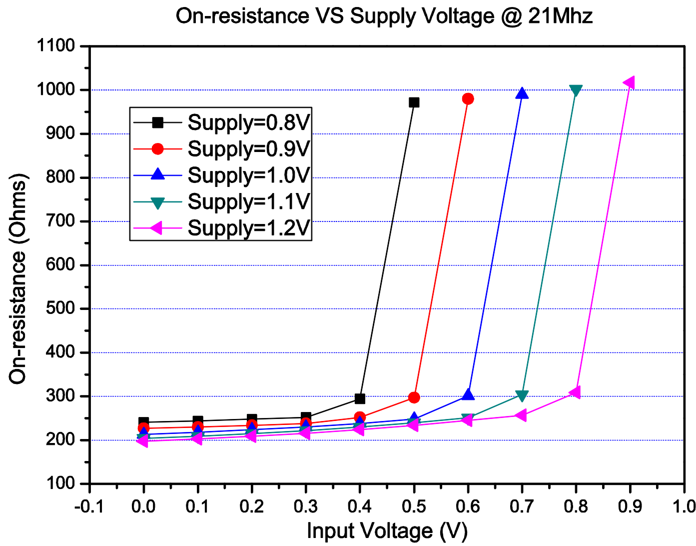


Fig.4.23 Plot of switch on-impedance vs supply voltage

Power spectral density simulations are demonstrated in Fig.4.22. The simulation is under a single tone input sine wave of 21MHz, the clock controlled sampling frequency is 150MHz. The 2nd order harmonic is basically limiting the SNDR performance without suppressing by differential topology. The SNDR is 59.85dBFS and SFDR is 76.02dBFS. To compare the power spectral density of the conventional one, of which SNDR is 57.73dBFS and SFDR is 73.14dBFS. This work can has a 2.12dBFS improvement in SNDR.

An on-resistance plot of the bootstrap switch is shown in Fig.4.23. It varies the supply voltage $\boldsymbol{V_{dd}}$ from 0.8V to 1.2V. The proposed bootstrapped switch shows low and reliable on impedance until $\boldsymbol{V_{in}}$ close to $\boldsymbol{V_{dd} - V_{thM8}}$.

4.4.4 LDE-aware LCR Effect on CMOS Switch Design

The top-plate sampling switch is implemented by utilizing an LDE-aware LCR. For a 11-bits, 5KS/s SAR ADC, the maximum on resistance of switch should be meets the charge time requirement. It can be estimated as:

$$\boldsymbol{R_{on} \times C_{total} \ll T_{per-conversion}},$$

$$\boldsymbol{T_{per-conversion} = \frac{1}{(N+2)f_{sample}}}. \quad \textbf{Eq. 4.19}$$

Where $\boldsymbol{C_{total}}$ is the capacitor array total capacitance, and $\boldsymbol{T_{per-conversion}}$ is the per-conversion time. N is the resolution bits. The maximum off resistance also should meet the accuracy requirement, that is:

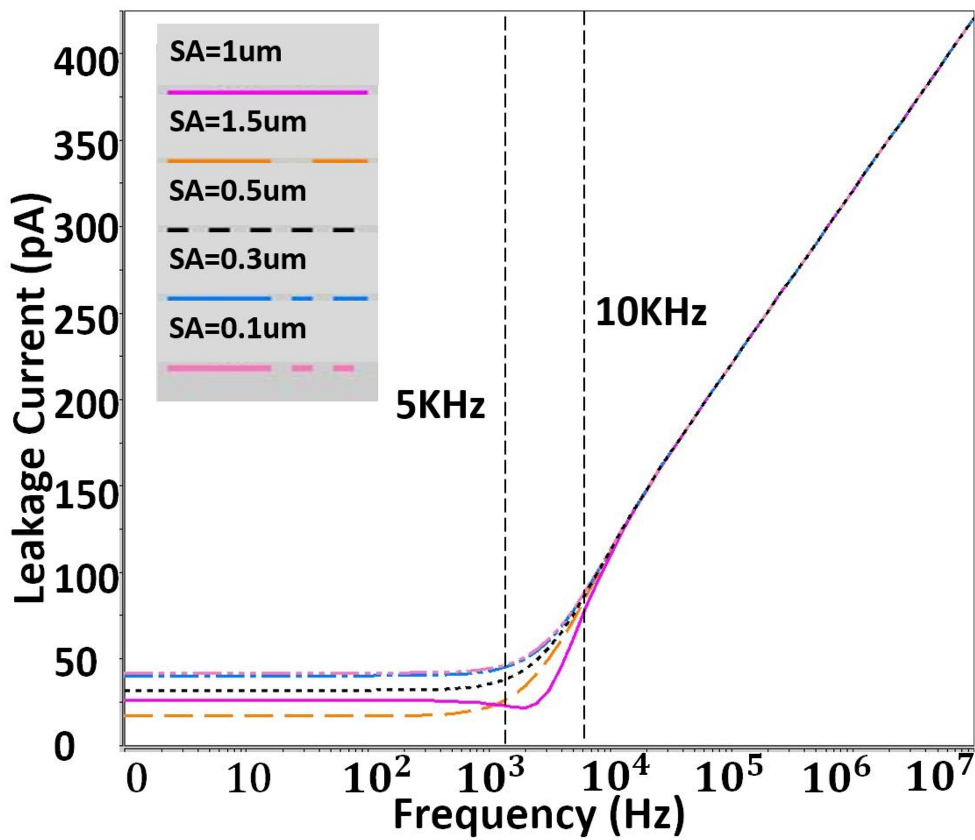
$$\boldsymbol{\frac{\triangle V}{T_{per-conversion}} = \frac{I_{leakage}}{C_{total}}}$$

$$\boldsymbol{\triangle V \ll 1\ LSB}. \quad \textbf{Eq. 4.20}$$

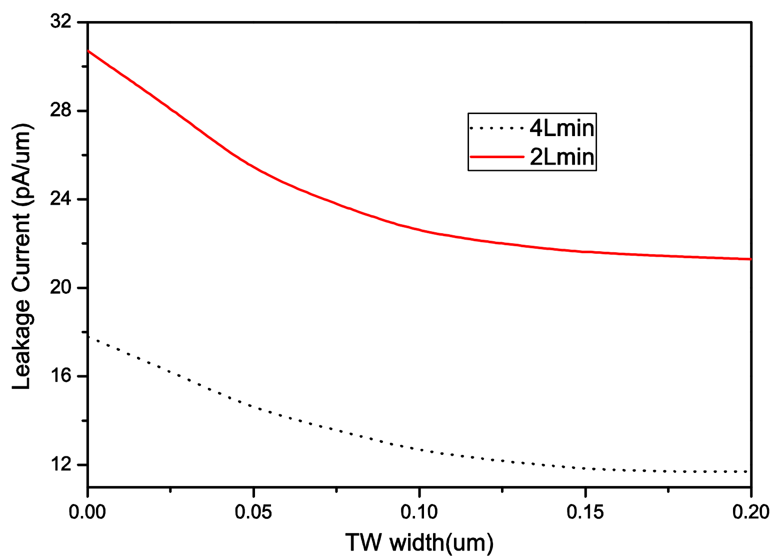
Besides, the 3dB bandwidth should meets the settling error requirements [70]

$$\boldsymbol{f_{3dB} > \frac{ln2(N+1)(N+2)}{\pi} f_{sample}}. \quad \textbf{Eq. 4.21}$$

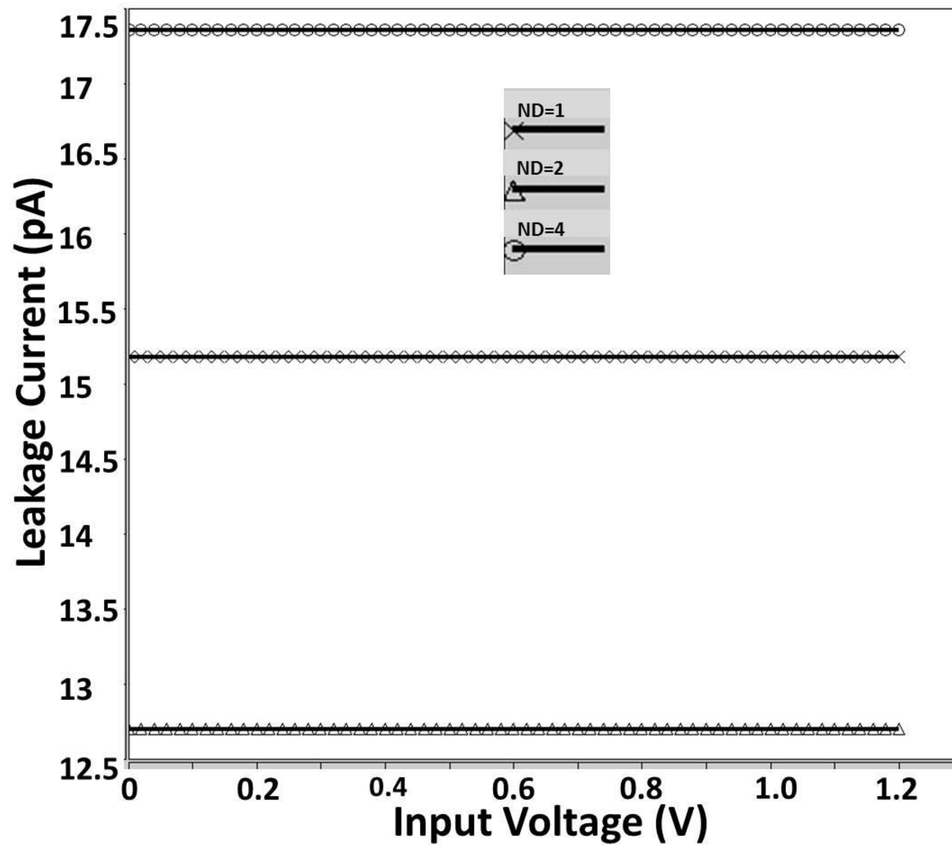
Based on Eq. (4.19), Eq. (4.20) and Eq. (4.21), the channel length can be roughly decided as $\boldsymbol{4L_{min}}$.



(a) Simulation of leakage current versus input frequency associating with different SA=SB length



(b) Simulation of leakage current versus TW width associating with different channel length



(c) Simulation of leakage current versus input voltage associating with ND numbers

Fig.4.24 Simulation Results in Switches

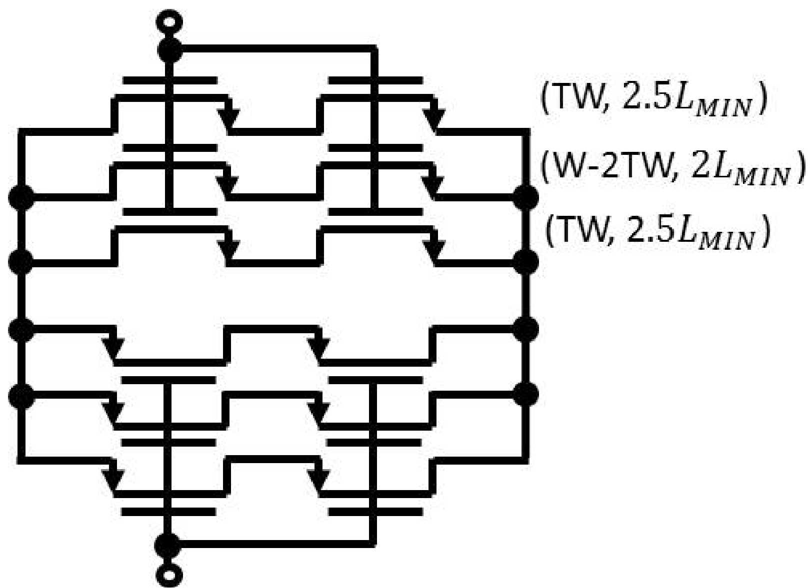


Fig.4.25 Switch MOSFETs equivalent circuit

For the reason that the mainstream of junction leakage flows through drain to source, the diffusion length plays a leading role of leakage value. As mentioned in before, the first step to decrease the leakage is to determine the SA=SB length. Fig.4.24 (a) shows HSPICE simulated leakage current versus input frequency associating with different SA=SB length. The junction leakage reduction is not significant when the input is over 10 KHz. When the input exceeds 1 KHz, the curve of SA=SB=1um has lower leakage current. Thus the final decision of SA=SB is 1um. According to Nyquist operation, Fig.4.24 (a) has further proved that the LDE-aware effect wanes with sampling rate over 20 KS/s.

Secondly, the channel shaping is implemented. Fig.4.24 (b) shows 3D TCAD simulation results by sweeping TW width. The TL length is $0.5\boldsymbol{L_{min}}$. At TW equals 0.1um, the leakage is 13pA/um. In the simulation, the leakage reduction is not significant as TW width over 0.1um. However, with the increasing of gate length, the dynamic performance becomes worse. Hence the TW width is set to 0.1 um in the layout. An observation of switches with

$2\boldsymbol{L_{min}}$ and $4\boldsymbol{L_{min}}$ channel length is also demonstrated in Fig. 8 (b). The longer channel length achieves better leakage performance.

The last step is to determine the finger numbers. Fig.4.24 (c) describes the HSPICE simulation of leakage current versus input voltage with different finger numbers. The results suggest a best leakage efficiency curve of ND=2.

Fig.4.25 shows the equivalent circuit of the final determined switch MOSFETs. It is combined with two edge transistors, which has a length of $2.\ 5\boldsymbol{L_{min}}$, width of TW. And the main transistor has length of $2\boldsymbol{L_{min}}$, width of W-2TW. The measurement results show the validity of LDE-aware LCR. The original DAC power is 29.3nW, after optimal design, the DAC power is 17.6nW. A totally 39.9% power dissipation is saved. Meanwhile, a worst-case SNDR at 5KS/s sampling, 2.41 KHz sine-wave input near-Nyquist operation is simulated in HSPICE, the results show SNR of original switch is 68.5dB, and the optimal switch is 68dB. The additional parasitic, which are caused by the LDE-aware LCR, are the main reason for the SNR degrading. The decreased SNR introduces a loss of 0.08-ENOB. However, the result still reasonable in this design requirement.

4.5 Implementation of LDE-aware LCR in comparator design

4.5.1 Design Basics

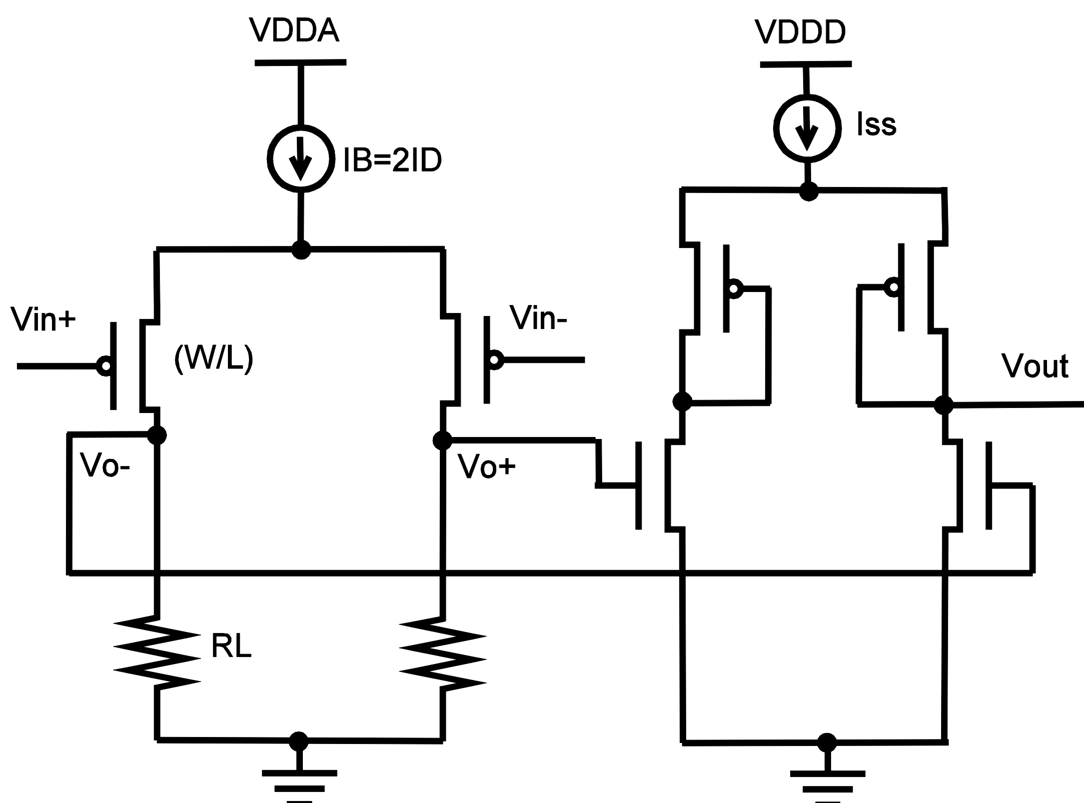


Fig.4.26 Basic comparator schematic

A comparator is the key building block as a 1 bit quantizer to resolve the sampled input V_{in} and the output of a DAC. The two main design factors in a comparator are thermal noise and offset. To meet the tight thermal noise and offset requirement in the high resolution SAR ADC, a preamplifier has been a good solution, but if the latch-only comparator without a preamplifier has sufficient margins for noise and offset, it is just a burden. The Fig.4.26 shows an example schematic of a pre-amplifier plus latch comparator. In addition, together a kick-back noise and the total capacitance of a DAC should be considered.

The offset of a comparator is the total offset of a SAR ADC which just limits little signal power but not linearity, and, therefore, it is out of considerations in most cases. However, note that when SAR ADCs are

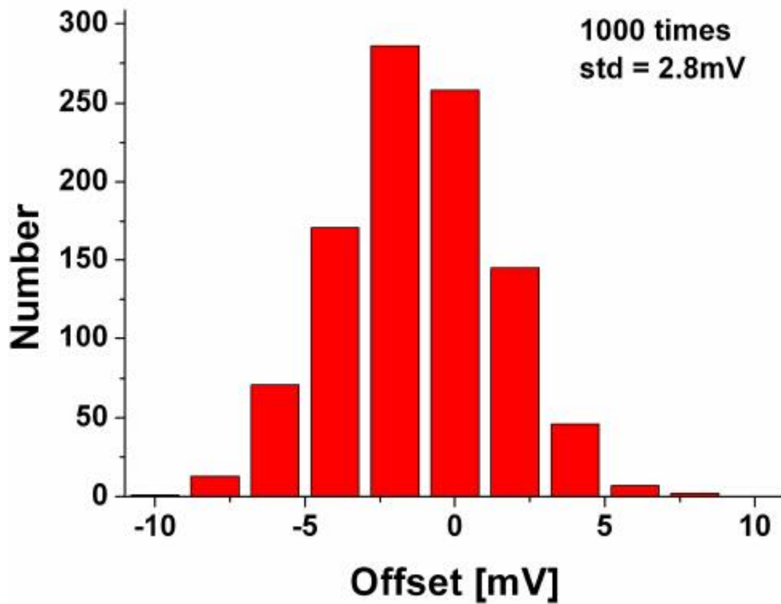


Fig.4.27 Offset simulation example

applied to time-interleaving technique or a comparator is used for a capacitor mismatch calibration, the offset should be calibrated Fig.4.27 shows an example of Monte-Carlo offset simulation to estimate the offset of a comparator. The 'std' means one sigma offset voltage

The thermal noise of a comparator is one of the main noise sources m a SAR ADC. In SAR ADCs, comparator thermal noise can limit the maximum achievable resolution. More than l and 2 ENOB reductions are observed in [59], [74] and [75], respectively, because of thermal noise, and degradations could be even worse with scaled supply voltages and the extensive use of dynamic regenerative latches without pre-amplification. Unlike mismatch, random noise cannot be compensated by calibration and would finally

demand a quadratic increase in power consumption unless alternative circuit techniques are devised.

4.5.2 Proposed comparator energy model

Fig.4.28 describes a normal transistor's capacitor model. Try to recall the LOD with variables of finger number (NF) introduce in Fig.2.1. SA and SB usually have the same size for symmetric arrangement. SD is set to as long as SA and SB. The poly gate length L is decomposed into 5 fingers, thus NF=5. If the two fingers in the ends of diffusion are set to dummy poly gates, then SA=SB=2SD+(1/5)L and NF=3.

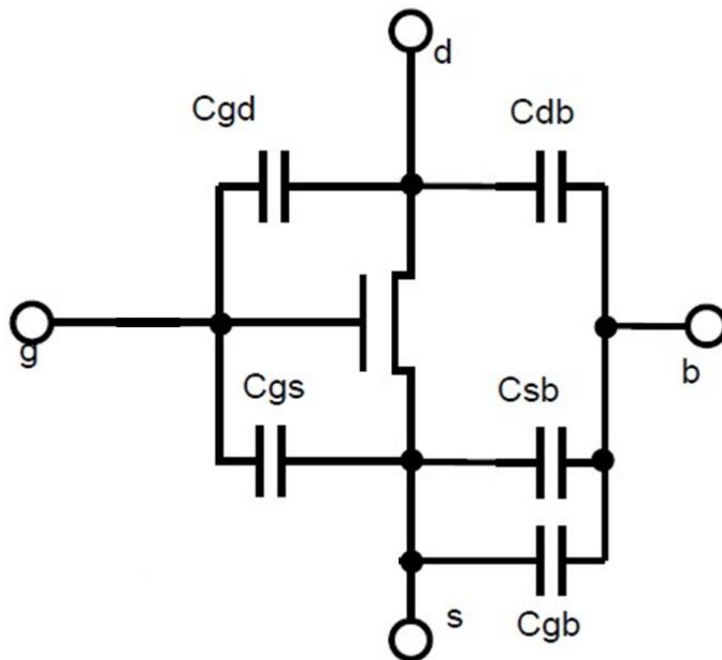


Fig.4.28 Transistor capacitor model

Using an approach similar to [76], we first note that the capacitive load can be a approximated by the amplifier's parasitic capacitance as:

$$C_L = C_{gslatch} + \sum_{i=1}^{n} C_{paraMn}. \qquad \textbf{Eq. 4.22}$$

Where $C_{gslatch}$ is the gate-source parasitic capacitance in the latch circuit input differential pairs, C_{paraMn} is the parasitic capacitance in source or drain of transistors. n is the number of latch related input transistors. The distributed gate-to-channel capacitance as seen between the gate and the source is approximated by (Transistors work in saturation region is assumed)

$$C_{gslatch} \cong 2/3 C_{OX} W_{latch} L_{latch}. \qquad \textbf{Eq. 4.23}$$

C_{paraMn} ban be approximated as:

$$C_{paraMn} \cong C_{diffusion} W \times SA, \qquad \textbf{Eq. 4.24}$$

where W_{latch} and L_{latch} is dimension of the latch circuit input differential pairs. In general, the transistor width required in the pre-amplifier is far greater than it is in the latch circuit. Thus we only consider the parasitic capacitors in the pre-amplifier as the input load of the latch circuit. Eq. (4.22) is simplified as:

$$C_L = C_{gslatch} + C_{diffusion} \sum_{i=1}^{n} W_{ampn} \times SA. \qquad \textbf{Eq. 4.25}$$

Where W_{ampn} is the width of the pre-amplifier input differential pairs. In Fig.4.27, n is 1.

The total SAR comparator energy per conversion, including
the contribution of the latch proposed in [77], is

$$E_{comp}^{SAR} = \frac{b+U}{f_S} I_B V_{DDA} + b C_{latch} V_{DDD}^2. \qquad \textbf{Eq. 4.26}$$

Here, b is resolution of the converter, in bits. U is number of periods for sampling in SAR ADC. f_S is sampling frequency. $I_B = 2I_D$. V_{DDA} and V_{DDD} are the analog and digital voltage supplies, respectively, and C_{latch}

represents the total switched capacitance in the latch, which scales in proportion with its input capacitance, so $\boldsymbol{C_{latch}} = \boldsymbol{C_L}$. Usually, increasing the latch offset to reduce comparator bias currents also reduces the latch energy.

A well-known saturation region drain current equation is

$$\boldsymbol{I_D = \frac{1}{2}\mu_{eff}C_{OX}\frac{W}{L}(V_{gs} - V_{th})^2} \qquad \textbf{Eq. 4.27}$$

Notice that the mechanical stress causes the electron mobility changing, the threshold voltage shifting and other second-order effects changing. Along with aggressively scaled-down CMOS technology, STI becomes one of the major factors determining circuit performance. In BSIM4 model, a STI stress model has already been developed which follow the 1/LOD trend but reveal different L and W scaling. We substitute Eq. (4.25), Eq. (2.2) and Eq. (4.27) into Eq. (4.26), the SAR comparator energy equation is rewritten as:

$$\boldsymbol{E_{comp}^{SAR} = \frac{b+U}{f_S}V_{DDA}\left(\frac{2K}{SA+0.5L}+1\right)\mu_{eff0} \times \left(V_{gs} - V_{th}\right)^2 \frac{W_{amp}}{L_{amp}}C_{OX} + bV_{DDD}^2\left(\frac{2}{3C_{OX}W_{latch}L_{latch}} + C_{diffusion}\sum_{i=1}^{n} W_{ampn} \times SA\right)}$$

Eq. 4.28

The only variable in Eq. (4.28) is SA. Thus we can see if we want to design a latch with smaller power consumption, we should minimized the SA subject to the design rules

4.5.3 LDE-aware LCR effect on comparator design

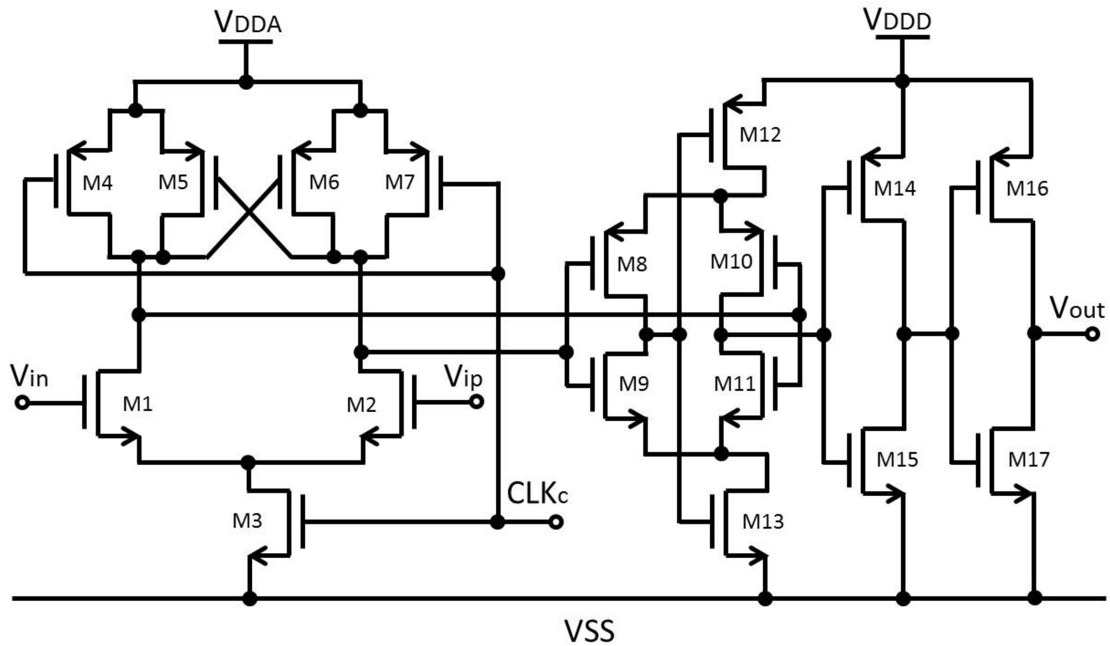


Fig.4.29 Dynamic latch comparator

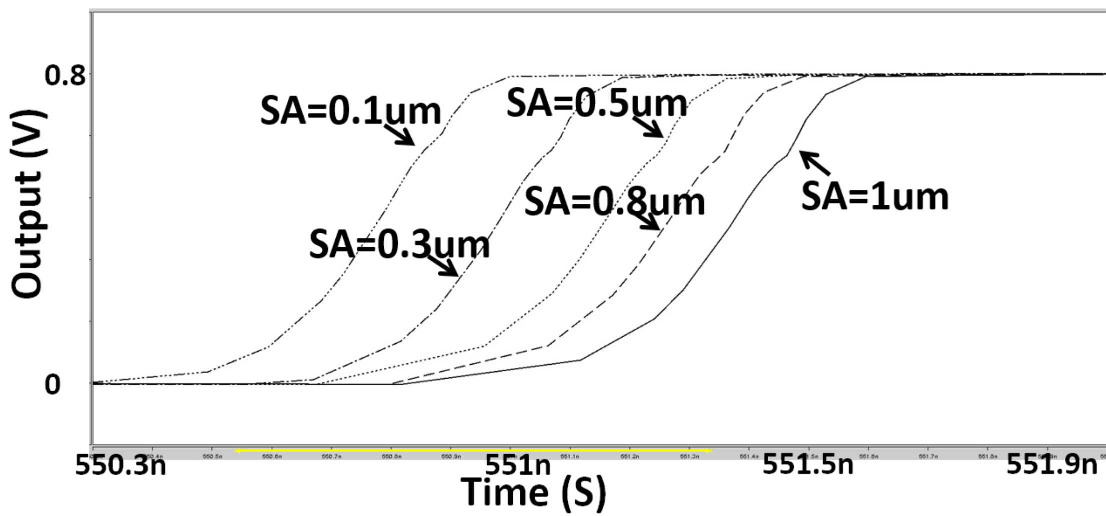


Fig.4.30 Setting time simulation result with different SA=SB length

The dynamic comparator is shown in Fig.4.29. The dynamic latch comparator has a 0.8V analog supply voltage ($\boldsymbol{V_{DDA}}$). The SR latch and two inverters with a 0.35V digital supply ($\boldsymbol{V_{DDD}}$) has been used to make the output loading identical.

Once again, the three steps of LDE-aware LCR are utilized in the comparator design. The producer is similar to the switch design, and it is not repeated here. The comparator performance should be invested and verified by employing the LDE-aware LCR. Normally, the comparator is designed to satisfy four specifications: offset, noise, gain and speed. Offset is very important because it contributes heavily to the whole linearity performance of the ADC. Since the input common-mode voltage is kept at mid-rail of the full-scale, the comparator offset appears as static offset [63].

$$\boldsymbol{V_{offset} = \triangle V_{th1,2} + \frac{\left(V_{gs} - V_{th}\right)_{1,2}}{2}\left(\frac{\triangle S_{1,2}}{S_{1,2}} + \frac{\triangle R_{1,2}}{R_{1,2}}\right). \mathrm{Eq.}\,4.29}$$

Monte Carlo simulations of the comparator offset showed a standard deviation offset of 5.1mV. After the LDE-aware LCR is applied, the standard deviation offset is 6.3mV. The diffusion parasitics caused the difference. Then we implemented LDE-aware LCR again with setting NF=1. The standard deviation offset is 5.3mV. This value is accepted in this work.

The primary comparator noise is thermal noise, which has the usual form as kT/C [78]. The post-layout extracted capacitor of the comparator is 6.6fF. After operating LDE-aware LCR (NF=1), the capacitance is 6.8fF. The comparator thermal noise is still in the same order as the quantization noise of the ADC.

For low-speed ADC, the gain and speed requirement is not so strict anymore. Fig. 11 shows the setting time simulation with different SA=SB

length (NF=1). For a microsecond level per-conversion time, the delay caused by LDE-aware LCR is accepted.

The measured power consumption results showed an original one of 1.6nW, while after LDE-aware LCR tuning, the value has changed to 0.9nw, totally 43.7% of comparator energy is saved.

4.6 Implementation of LDE-aware LCR in SAR control logic

4.6.1 Synchronous and Asynchronous SAR controller

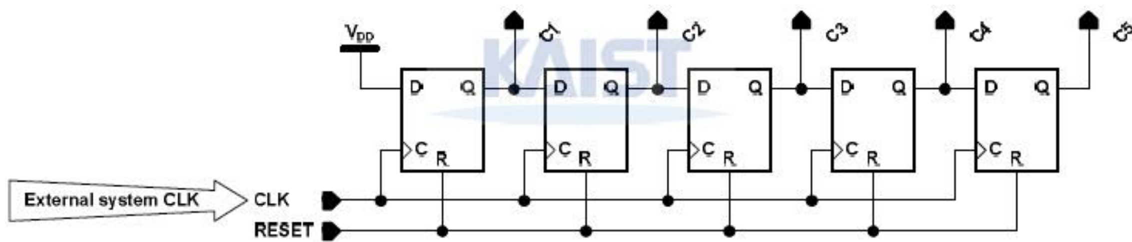


Fig.4.31 Synchronous 5 bit SAR control logic

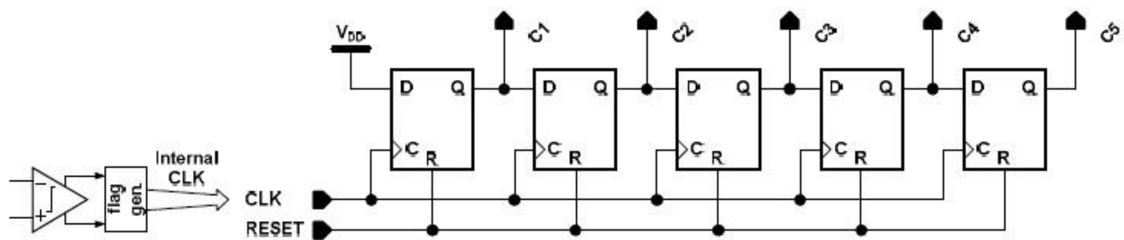


Fig.4.32 Asynchronous 5 bit SAR control logic

Before 2006, the controller for most SAR ADCs is designed synchronously. This means as illustrated in Fig.4.31 that overall control signals are generated by an external system CLK which is usually '1+resolution' times faster than an assigned sampling frequency (fs). All output codes from MSB to LSB is be resolved one by one. After 5 times conversions, all DFFs are reset by the rising edge of a system CLK only for the logic high period of C5. The synchronous SAR controller is driven by a high frequency CLK driver.

The Fig.4.32 shows the schematic of an asynchronous 5 bit SAR controller. A DFF array is identical to a synchronous one. Only one difference is the sources of CLK, which are the two outputs of a comparator. Exclusive-OR (XOR) gate driven by comparator performs the roll of flag generator, whose output will be high when the comparator has two different outputs in the latching phase, and whose output will be low when the two outputs of a comparator are the same in the reset phase as if the "flag" is an external high frequency CLK. In an asynchronous decision, the signal "flag" detected with XOR indicates the completing a comparison. The rising edge and falling edge of this CLK shows up asynchronously by irregular repetitions of latching and reset of a comparator, because the comparator's latching time is logarithmically proportional to its input magnitude [56]. The SAR ADC with asynchronous controller can increase the conversion speed effectively, because the DAC settling operation is started as soon as the flag goes up. This is different with a synchronous SAR ADC which must unconditionally wait the next falling edge of a system CLK to start DAC settling after comparison at the rising edge. The ratio of total latching time of asynchronous and synchronous SAR ADC ($\boldsymbol{T}_{\mathbf{asynh}}$ /$\boldsymbol{T}_{\mathbf{synh}}$) approaches 1/2 as the resolution increases. Besides, this controller does not require external high frequency system CLK.

4.6.2 Proposed SAR control logic with LDE-aware LCR

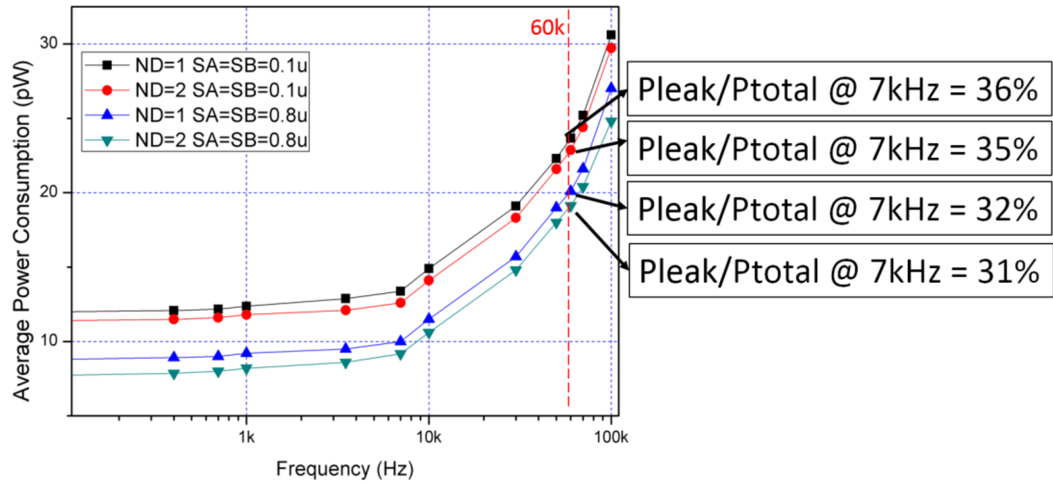


Fig.4.33 Simulated average power consumption versus operating frequency of SAR control logic

The operating frequency of the SAR logic is 60 kHz, the leakage power plays an important role of the total power. LDE-aware LCR also applied to reduce the leakage power, including set SA=SB=0.8um, NF=2, increased the channel length to 2*Lmin*. The gate channel shaping could not use because that the digital standard cells already follow the minimum width concept. Fig. 12 demonstrate the average power consumption of SAR logic control as a

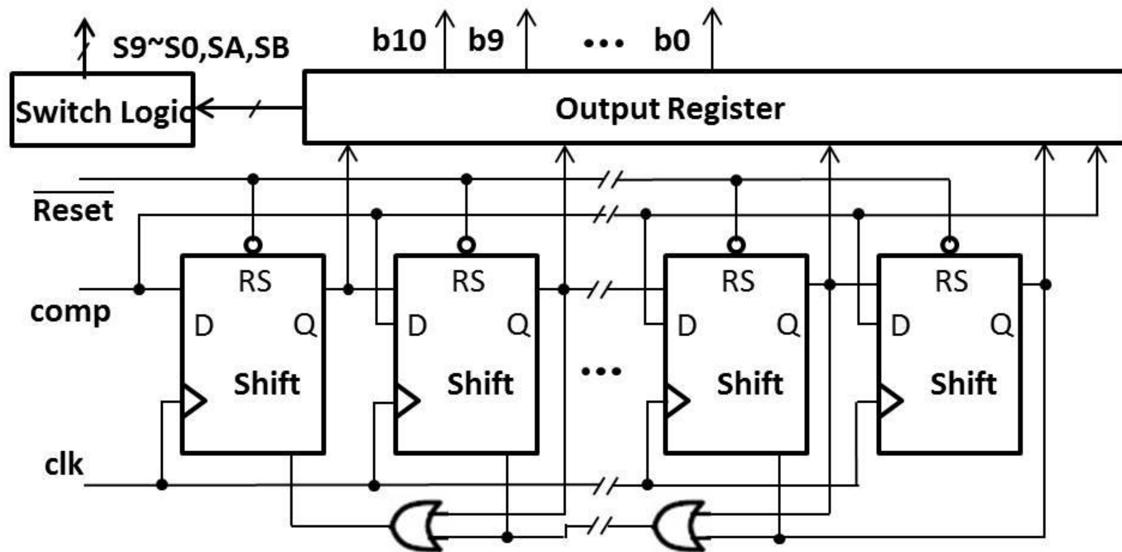


Fig.4.34 SAR control logic schematic

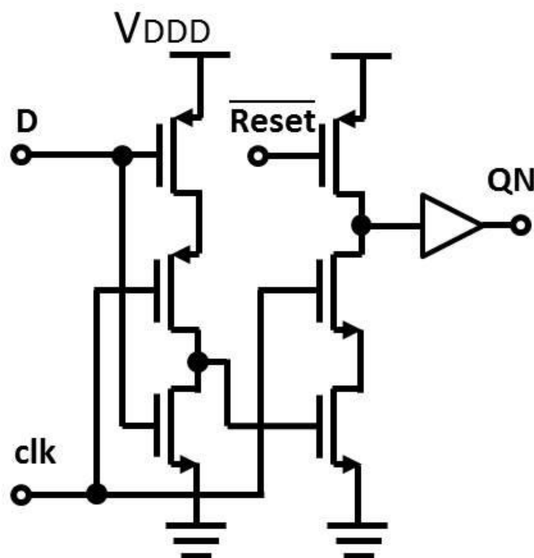


Fig.4.35 Once-triggered D flip-flop

function of its clock frequency. In the operating frequency of $clk = 60\ kHz$, the leakage power can constitute at least 31%of the total power.

The digital logic control utilizes a synchronous phase generator configuration, shown in Fig.4.34. Asynchronous process [63], [68] has been frequently used for high-speed SAR ADCs in order to avoid using a high-

frequency clock generator [79]. However, in this work, the low sampling rate should not require such high-speed. Therefore, a synchronous SAR logic is much more proper. In the sampling mode, the ’Reset’ control signal goes low and all the outputs of shift register except the one determined b10 are cleared. The b10 decisive shift register accepts the first comparison result by means of the clock triggering, in which both the OR logic gate’s input and b10 is determined. The next, b10 is latched in the output register. In the meantime, the input of the second shift register transmits to the output of b9. This procedure repeats in the same way. Since there is no capacitor needed to switch after the final comparison, the register for LSB can be saved.

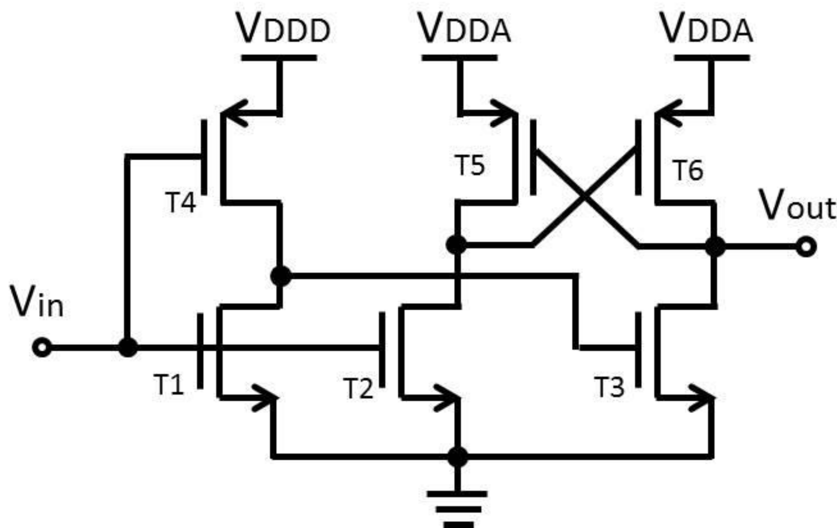


Fig.4.36 Level shifter

For low-power implementation, a once-triggered D flip-flop (DFF) [71] is introduced. Fig. 14 shows the circuit structure. Compared with master-slave DFF, the once-triggered DFF has eight transistors, which reduce the power consumption, gate delay and active area.

A significant digital power reduction technology is reported in [70]. A low digital supply voltage of 0.35V has been employed. The logic control outputs are transferred by level shifter, shown in Fig. 15. A simpler scheme

is adopted to instead of the original circuit. The total level shifter has six transistors, which behaves lower power dissipation.

4.7 Capacitor array considerations

In the SAR ADC design, we use the metal-oxide-metal (MOM) capacitors to construct the capacitor array. According to the analysis in [66], the unit capacitance with the random variation ϵ_C causes increasing the value of DNL at the critical transition. Therefore, for high yield, it is necessary to maintain a standard deviation as:

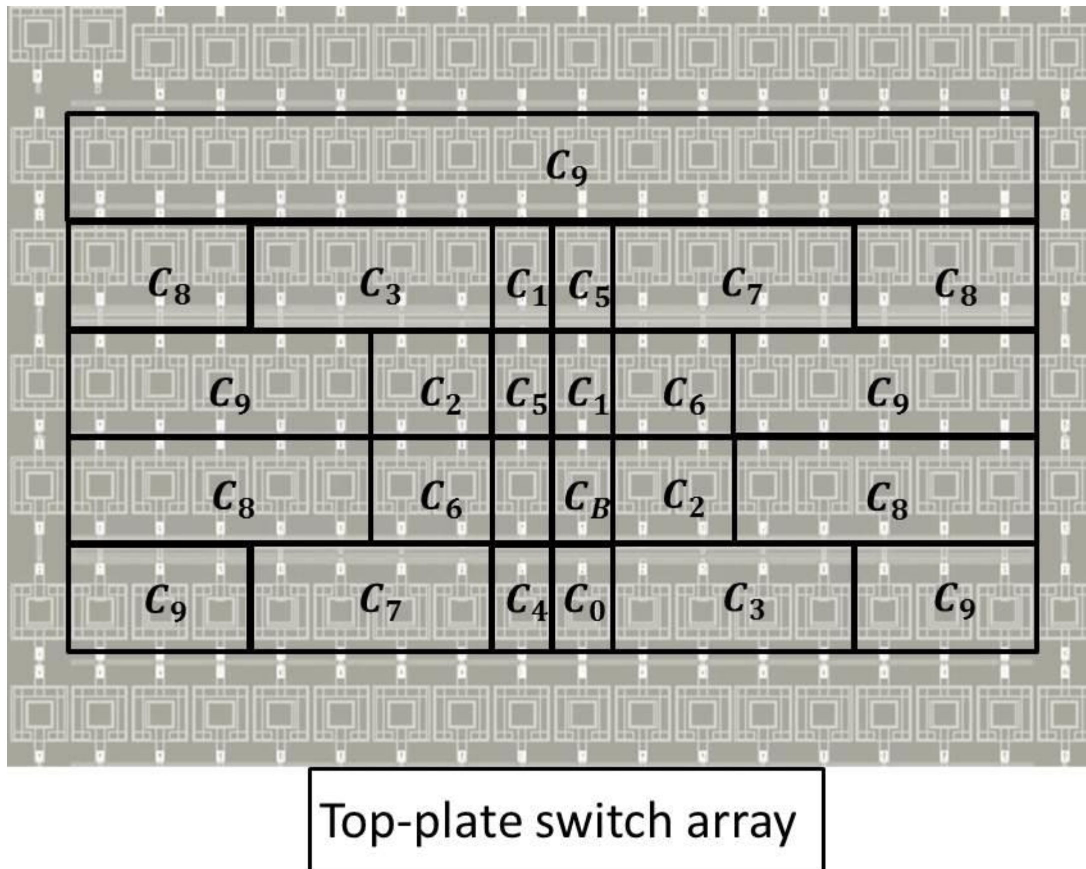


Fig.4.37 Layout of the split-capacitor array which follows a switch oriented common-centroid configuration. The capacitors are indicated according to Fig.4.14

$$3\sigma_{DNL,MAX} = 3 \times \sqrt{2}\epsilon_C \times 2^{(N-1)/2} LSB \leq \frac{1}{2} LSB. \text{Eq. } 4.30$$

In this technology, the MOM capacitor has a density of $1fF/um^2$, and matching of 0.433%/um. It leads to a minimum unit capacitance of 2.95fF. For the split-capacitor array, the unitary capacitor should be at least 15fF due to the more stringent matching requirement. The two capacitor networks occupy a total active area of 97um×255um, about 68% of the whole ADC.

Besides the mismatch associated with capacitor sizing, the crossing parasitics exist in routing wires also comparatively influences the linearity performance. Fig.4.37 illustrates the layout floor plan for the capacitor array. The unitary capacitors of which determines the lowest bits at MSB section and LSB section ($C4$ and $C0$), and the unbinary weighted capacitor (CB) were placed to close the top-plate switch array. The switch oriented common-centroid configuration is helpful to reduce the parasitic effect on the lowest bits, which are easily leading larger numerical point in static measurement (DNL/INL).

4.8 Results

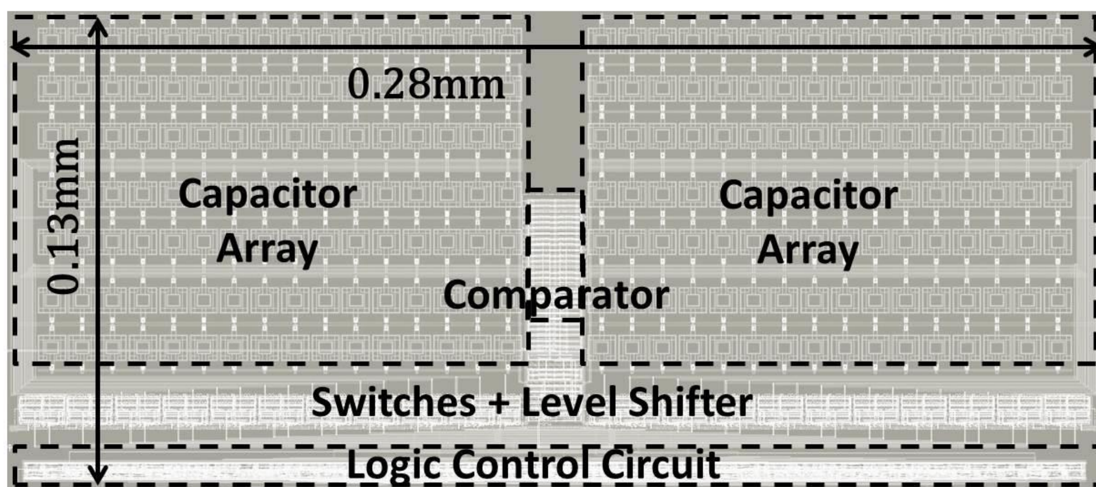


Fig.4.38 SAR ADC layout

LDE-aware LCR	Original	Improving
Comparator	1.6	0.9
Capacitive DAC	29.3	17.6
Control Logic	24.3	20.1
Total	55.2	38.6

Table4.2 POWER CONSUMPTION OF SAR ADC COMPONENTS (nW)

The proposed SAR ADCs are fabricated in a triple-well 65nm/1.2V CMOS technology. Fig.4.38 shows the layout of the proposed SAR ADC. The active area is 130 $\times 279\boldsymbol{um^2}$. Besides the capacitor array, the comparator occupies 1.3% area. The switch array occupies 6.8% area. The level shifter accounts for 1.9% (which has a total number of twelve). The logic control circuit occupies 7.4% area. The rest area is used for routing, and the spare area is filled with metal ground. The wafer is packaged with bonding wires. The test

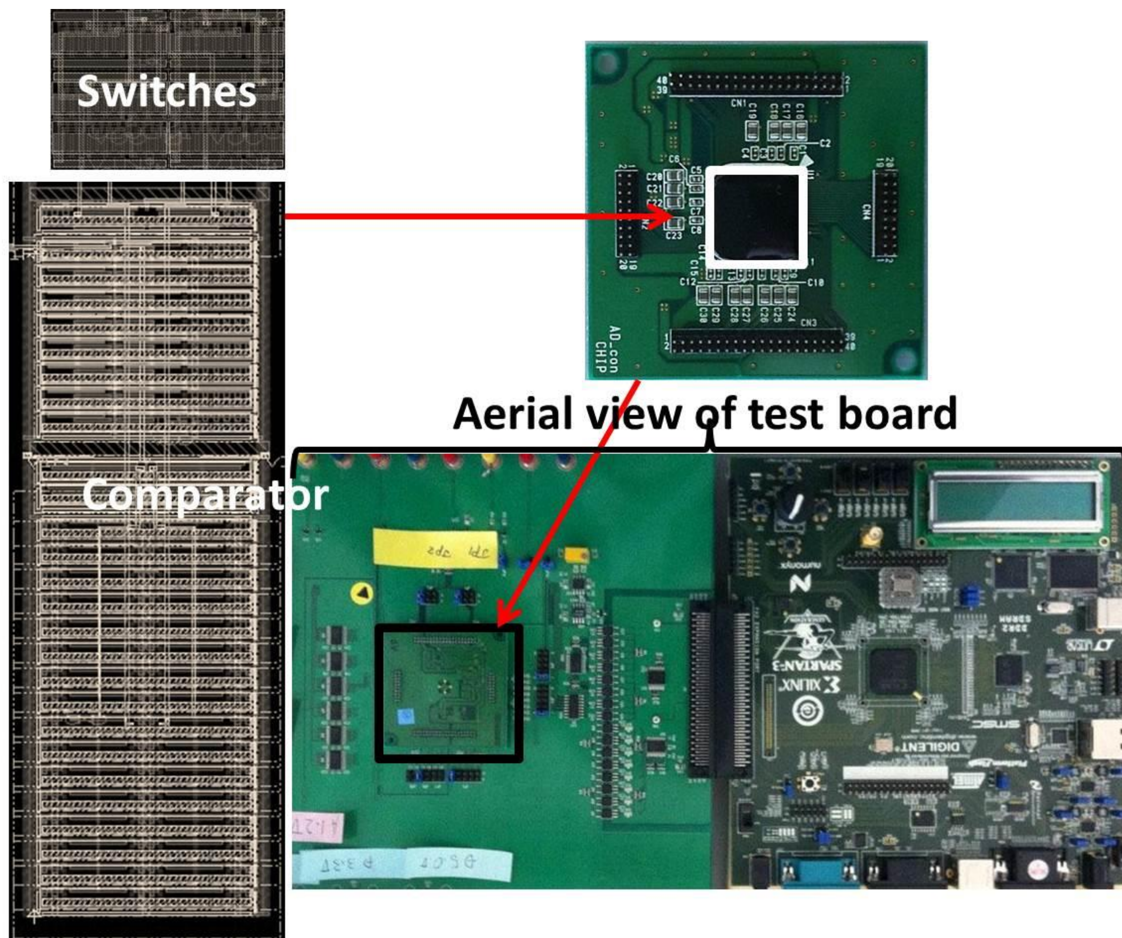
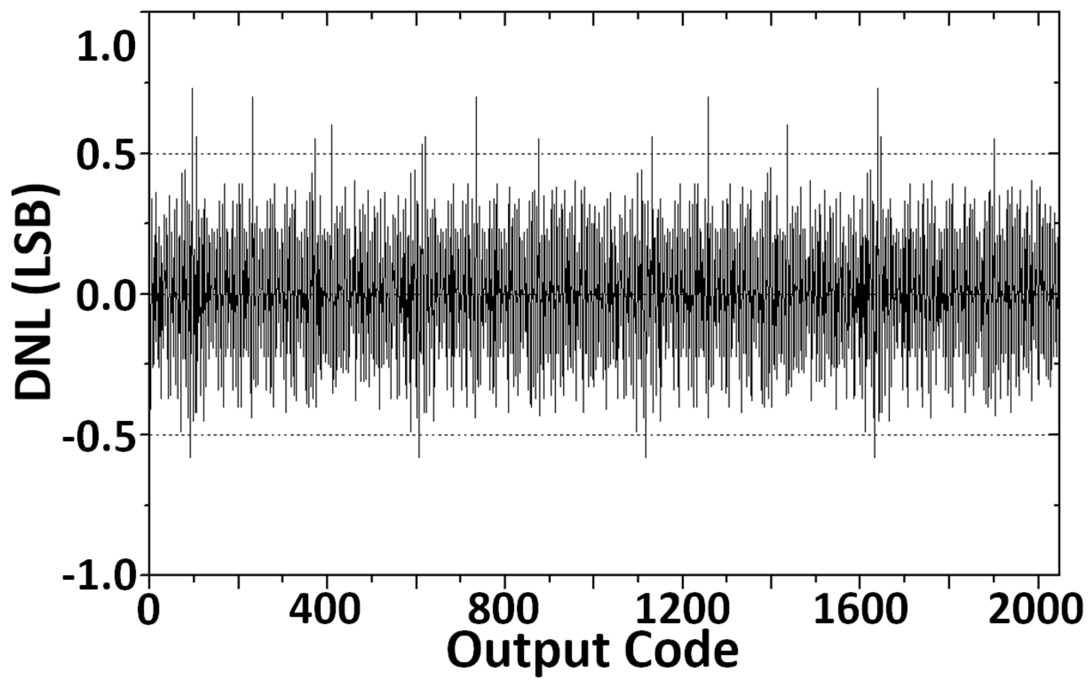


Fig.4.39 SAR ADC layout test board with the same analog scheme

Technology	65-nm CMOS	
Resolution	11-bit	
Sampling Rate	5-KS/s	
Supply Voltage	0.8-V/0.35-V	
Full Scale Analog Input	1.67-V_{pp} differential	
LSB	1.172-mV	
LDE-aware LCR	with	w/o
DNL	+0.73/-0.62 LSB	+0.7/-0.61 LSB
INL	+0.37/-0.83 LSB	+0.37/-0.79 LSB
SNDR (f_{in} 2.4KHz) SFDR ($f_{sampling}$ 5KHz)	55.93 dB 64.47 dB	57.07 dB 66.1 dB
ENOB (bits)	9	9.19
$FOM = \frac{POWER}{2^{ENOB} \times min\{f_s, 2 \times ERBW\}}$ (Unit is fJ/conversion-step)	15.1	18.9

Table4.3 SUMMARY OF PERFORMANCE



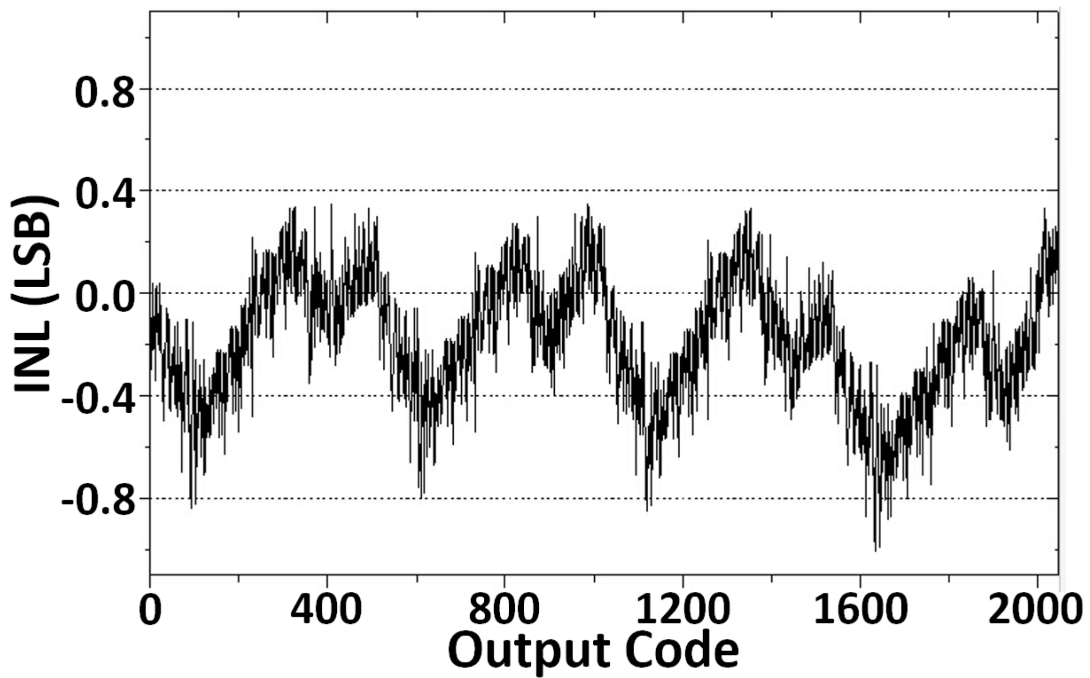


Fig.4.40 Simulated DNL and INL

board in Fig.4.39 is composed of a plug-in board (in the left side of the test board aerial view) and a Xilinx FPGA digital control board (in the right side).

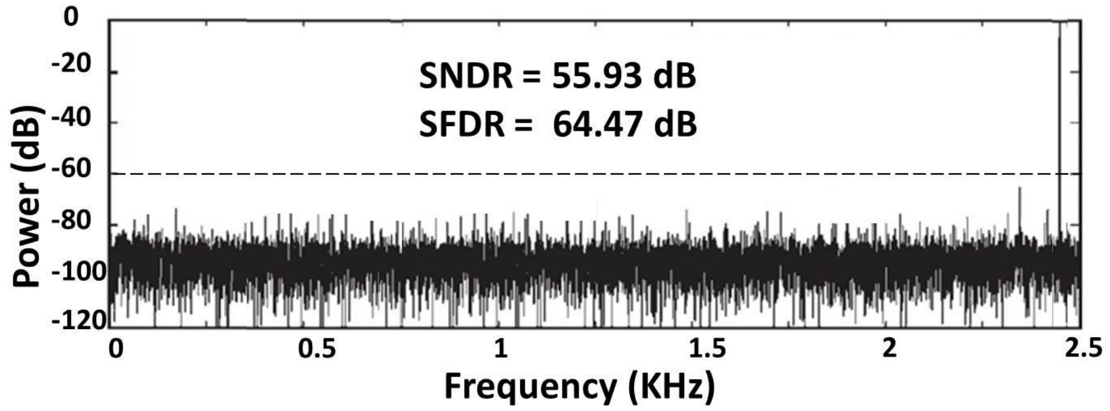


Fig.4.41 Simulated 8192-point FFT spectrum at 5 KHz

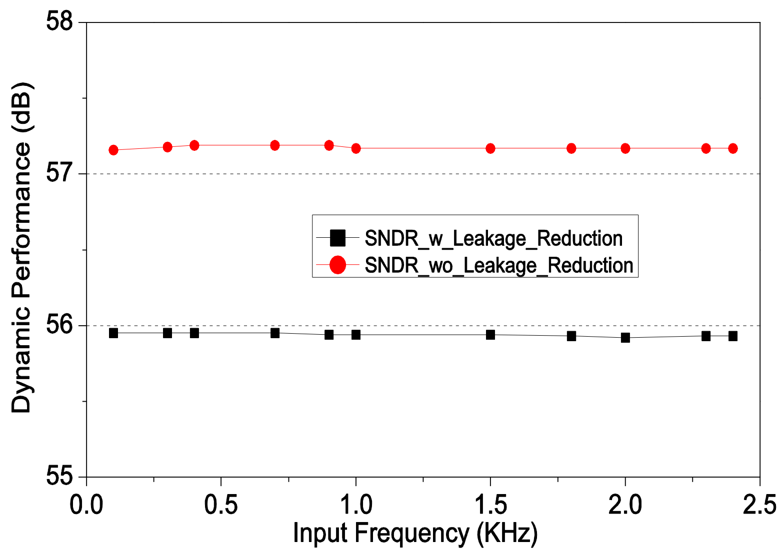


Fig.4.42 Dynamic performance comparison versus input frequency

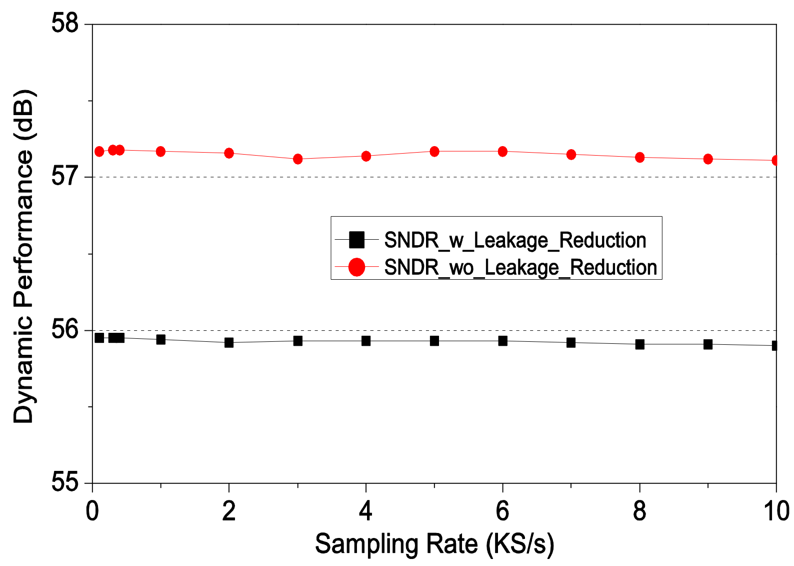


Fig.4.43 Dynamic performance comparison versus sampling frequency

Reference	ISSCC'12	CICC'11	APCCAS'12	ISSCC'11	ISSCC'13	This Work (Post-layout Results)
Architecture	SAR	SAR	SAR	SAR	SAR	SAR
Technology	0.13μm	0.18μm	0.18μm	65nm	65nm	65nm
Resolution (bit)	10	10	9	10	12	11
Sampling Rate (KS/s)	1	1	1	20	40	5
Supply Voltage(V)	1(analog)/0.4(digital)	0.5	0.55	0.55	0.6	0.8(analog)/0.35(digital)
Differential Input Range	NA	0.5V	NA	NA	NA	1.3V
SNDR (dB)	56.7	53.05	54.2	55	62.5	55.93
ENOB (bits)	9.1	8.52	8.7	8.84	10.1	9
Power (nW)	53	2.5	9.87	206	97	38.6
FOM (fJ/conversion-step)	94.5	6.8	23.7	22.4	2.2	15.1
Active Area (mm^2)	0.19	0.05	0.04	0.212	0.076	0.036

Table 5 COMPARISON TO STATE-OF-THE-ART WORKS

The measurement power consumption results confirm the validation of LDE-aware LCR. Empirically, the rest SAR ADC performances from post-layout simulations are close to the measured results.

4.8.1 Measured Power Consumption Results

For the purpose of verifying the feasibility of the proposed LDE-aware LCR. The power consumption of ADC is estimated from the measured results. Both SAR ADCs with LDE-aware LCR and without LDE-aware LCR are taped out. SAR ADC with LDE-aware LCR consumes 38.6-nW, while without LDE-aware LCR dissipates 55.2-nW. The partition of the power consumption is listed in Table4.2. Comparing with the every portion, the energy saving is listed as 44% of comparator, 40% of capacitive DAC and 17% of logic control, respectively. From the experience of the preliminary work, the power consumption result from post-layout is close to the measured one. LDE-aware LCR is much more effective in analog portion because of the higher operating frequency (60 kHz) in digital standard cells. Totally, about 30% of leakage power is eliminated.

4.8.2 Post-Layout Simulation Results of SAR ADC Performance

The rest of SAR ADC performances are based on the post layout simulations. The simulated DNL and INL are shown in Fig.4.40. The maximum DNL is +0.73/-0.62LSB. The maximum INL is +0.37LSB/-0.79LSB. The SNDR and SFDR of the ADC is simulated by using single tone testing. A fast Fourier transform (FFT) spectrum with an input frequency of 2.481 KHz at a 5KS/s sampling rate is shown in Fig.4.41. The simulated SNDR and SFDR are 55.93dB and 64.47dB, respectively. Fig.4.42 plots the simulation result

of the dynamic performance versus the input frequency at 5KS/s sampling rate. Fig.4.43 plots the dynamic performance versus the sampling frequency with a sinusoidal input at 2.1 KHz. The resultant ENOBs are 9-bits.

The performance of the ADCs are summarized in Table4.3. The LDE-aware LCR introduce additional parasitics because of the wider size tuning in diffusion region. In general, the performance of the proposed SAR ADC drops weakly. But the loss is in the tolerated range of the initial requirements. Comparing with the figure of merits (FOMs) between these two ADCs, FOMs are being improved 3.8-fJ/conversion-step with LDE-aware LCR. In Table 4.4, we show the ADC FOM performance under different sampling rate. The ADC achieves the optimal performance at 5Ks/s with the lowest FOM of 15.1-fJ/conversion-step. Actually, leakage current in a few nano-watt level power consumption ADCs (at 1KS/s), has a large dependency on the process (mentioned in Table2.1, parameters can be such as S/D depth, S/D doping etc.), supply voltage, and temperature. Even if the rest of static and dynamic performances are post layout simulated, the proposed power efficiency LDE-aware LCR has been verified practically without process limitations. Furthermore, the power efficiency of the proposed ADC is reasonable to be employed in medical implant devices. In Table 5, we compare the proposed ADC with other state-of-the-art ADCs which are reported recently [70], [80]-[83]. The proposed ADC achieves the minimize layout dimensions.

References

[52] B.Murmann, “ADC Performance Survey 1997-2013,” http://www/stanford.edu/~murmann/adcsurvey.html.

[53] Suarez, R.; Gray, P.; Hodges, D., "An all-MOS charge-redistribution A/D conversion technique," Solid-State Circuits Conference. Digest of Technical Papers. 1974 IEEE International, vol.XVII, no., pp.194,195, 15-13 Feb. 1974

[54] Walden, R.H., "Analog-to-digital converter survey and analysis," Selected Areas in Communications, IEEE Journal on , vol.17, no.4, pp.539,550, Apr 1999

[55] R.Schreier and G.C.Temes, “Understanding Delta-Sigma Data Converters”. New York; John Wiley & Sons, 1ed., 2005.

[56] Chen, S.-W.M.; Brodersen, R.W., "A 6-bit 600-MS/s 5.3-mW Asynchronous ADC in 0.13- μm CMOS," Solid-State Circuits, IEEE Journal of , vol.41, no.12, pp.2669,2680, Dec. 2006

[57] Wenbo Liu; Yuchun Chang; Szu-Kang Hsien; Bo-Wei Chen; Yung-Pin Lee; Wen-Tsao Chen; Tzu-Yi Yang; Gin-Kou Ma; Yun Chiu, "A 600MS/s 30mW 0.13µm CMOS ADC array achieving over 60dB SFDR with adaptive digital equalization," Solid-State Circuits Conference - Digest of Technical Papers, 2009. ISSCC 2009. IEEE International , vol., no., pp.82,83,83a, 8-12 Feb. 2009

[58] Wenbo Liu; Pingli Huang; Yun Chiu, "A 12b 22.5/45MS/s 3.0mW 0.059mm^2 CMOS SAR ADC achieving over 90dB SFDR," Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International , vol., no., pp.380,381, 7-11 Feb. 2010

[59] Kuttner, F., "A 1.2V 10b 20MSample/s non-binary successive approximation ADC in 0.13/spl mu/m CMOS," *Solid-State Circuits Conference, 2002. Digest of Technical Papers. ISSCC. 2002 IEEE International* , vol.1, no., pp.176,177 vol.1, 7-7 Feb. 2002

[60] Chi-Hung Lin; Bult, K., "A 10-b, 500-MSample/s CMOS DAC in 0.6 mm^2," Solid-State Circuits, IEEE Journal of , vol.33, no.12, pp.1948,1958, Dec 1998.

[61] M. Yoshioka, K. Ishikawa, T. Takayama, and S. Tsukamoto, ”A 10 b 50 MS/s 820 W SAR ADC with on-chip digital calibration”, in IEEE ISSCC Dig. Tech. Papers, 2010, pp. 384-385.

[62] W. Y. Pang, C. S. Wang, Y. K. Chang, N. K. Chou, and C. K. Wang, ”A 10-bit 500-KS/s low power SAR ADC with splitting capacitor for bio-medical applications”, in Proc. IEEE A-SSCC, 2009, pp. 149-152.
[63] C. C. Liu, S. J. Chang, G. Y. Huang, and Y. Z. Lin, ”A 10-bit 50-MS/s SAR ADC with a monotonic capacitor switching procedure,” IEEE J. Solid-State Circuits, vol. 45, pp. 731-740, Apr. 2010.
[64] Guerber, J.; Venkatram, H.; Taehwan Oh; Un-Ku Moon, ”Enhanced SAR ADC energy efficiency from the early reset merged capacitor switching algorithm,” Circuits and Systems (ISCAS), 2012 IEEE International Symposium on , vol., no., pp.2361-2364, 20-23 May, 2012
[65] Y. Zhu, C.-H. Chan, U-F. Chio, S.-W. Sin, S.-P. U, R. P. Martins, and F. Maloberti, ”A 10-bit 100-MS/s reference-free SAR ADC in 90 nm CMOS,” IEEE J. Solid-State Circuits, vol. 45, pp. 1111-1121, Jun. 2010.
[66] B.P.Ginsburg and A.P.Chandrakasan, ”500-MS/s 5-bit ADC in 65-nm CMOS with split capacitor array DAC,” IEEE TJ.Solid-State Circuits, vol. 42, no. 4, pp. 739-747, Apr. 2007.
[67] V. Hariprasath, J. Guerber, S.-H. Lee, U. Moon, ”Merged capacitor switching based SAR ADC with highest switching energy-efficnecy,” Electron. Lett., vol. 46, pp. 620-621, Apr. 2010.
[68] Agnes, A.; Bonizzoni, E.; Malcovati, P.; Maloberti, F., ”A 9.4-ENOB 1V 3.8?W100kS/s SAR ADC with Time-Domain Comparator,” Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International , vol., no., pp.246,610, 3-7 Feb. 2008
[69] Ruoyu Xu; Bing Liu; Jie Yuan, ”Digitally Calibrated 768-kS/s 10-b Minimum-Size SAR ADC Array With Dithering,” Solid-State Circuits, IEEE Journal of , vol.47, no.9, pp.2129-2140, Sept. 2012
[70] Zhang, Dai ; Bhide, Ameya K. ; Alvandpour, Atila, ”A 53-nW 9.1-ENOB 1-kS/s SAR ADC in 0.13-μm CMOS for Medical Implant Devices,” Solid-State Circuits, IEEE Journal of , vol. 47, issue:7, pp. 1585-1593, July 2012.

[71] Huang, Guan-Ying ; Chang, Soon-Jyh ; Liu, Chun-Cheng ; Lin, Ying-Zu, ”A 1-μ*W* 10-bit 200-kS/s SAR ADC With a Bypass Window for Biomedical Applications,” Solid-State Circuits, IEEE Journal of , vol. 47, issue 11, pp. 2783- 2795, Nov. 2012.
[72] A. M. Abo and P. R. Gray, A 1.5-V, 10-bit, 14.3-MS/s CMOS pipeline analog-to-digital converter, IEEE J. Solid-State Circuits, vol. 34, no. 5, pp. 599606, May 1999.
[73] Gu[73]an-Ying Huang; Soon-Jyh Chang; Chun-Cheng Liu; Ying-Zu Lin, ”A 1-W 10-bit 200-kS/s SAR ADC With a Bypass Window for Biomedical Applications,” Solid-State Circuits, IEEE Journal of , vol.47, no.11, pp.2783,2795, Nov. 2012
[74] Craninckx, J.; Van der Plas, G., "A 65fJ/Conversion-Step 0-to-50MS/s 0-to-0.7mW 9b Charge-Sharing SAR ADC in 90nm Digital CMOS," *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International* , vol., no., pp.246,600, 11-15 Feb. 2007
[75] Giannini, V.; Nuzzo, P.; Chironi, V.; Baschirotto, A; Van der Plas, G.; Craninckx, J., "An 820μW 9b 40MS/s Noise-Tolerant Dynamic-SAR ADC in 90nm Digital CMOS," *Solid-State Circuits Conference, 2008. ISSCC 2008. Digest of Technical Papers. IEEE International* , vol., no., pp.238,610, 3-7 Feb. 2008
[76] E. Crain and M. H. Perrott, ”A numerical design approach for high speed differential, resistor-loaded CMOS amplifiers,” in Proc. IEEE Int. Symp. Circuits and Systems (ISCAS), 2004, pp. V-508-V-511.
[77] B. P. Ginsburg and A. P. Chandrakasan, ”Dual time-interleaved successive approximation register ADCs for an ultra-wideband receiver,” IEEE J. Solid-State Circuits, vol. 42, no. 2, pp. 247-257, Feb. 2007.
[78] Jaeha Kim; Leibowitz, B.S.; Jihong Ren; Madden, C.J., ”Simulation and Analysis of Random Decision Errors in Clocked Comparators,” Circuits and Systems I: Regular Papers, IEEE Transactions on , vol.56, no.8, pp.1844,1857, Aug. 2009 doi:
[79] M. Van Elzakker, E. Van Tujil, P. Geraedts, D. Schinkel, E. A. M. Klumperink, and B. Nauta, ”A 10-bit charge-redistribution ADC consuming 1.9 W at 1ms/s,” IEEE J. Solid-State Circuits, vol. 45, no. 5, pp. 10071015, May 2010.
[80] Tsung-Che Lu; Lan-Da Van; Chi-Sheng Lin; Chun-Ming Huang, ”A 0.5V 1KS/s 2.5nW8.52-ENOB 6.8fJ/conversion-step SAR ADC for biomedical applications,”

Custom Integrated Circuits Conference (CICC), 2011 IEEE , vol., no., pp.1,4, 19-21 Sept. 2011
[81] Anh Tuan Do; Chun Kit Lam; Yung Sern Tan; Kiat Seng Yeo; Jia Hao Cheong; Lei Yao; Meng Tong Tan; Minkyu Je, ”A 9.87 nW 1 kS/s 8.7 ENOB SAR ADC for implantable epileptic seizure detection microsystems,” Circuits and Systems (APCCAS), 2012 IEEE Asia Pacific Conference on , vol., no., pp.1,4, 2-5 Dec. 2012
[82] Yip, M.; Chandrakasan, A.P., ”A resolution-reconfigurable 5-to-10b 0.4-to-1V power scalable SAR ADC,” Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International , vol., no., pp.190,192, 20-24 Feb. 2011
[83] Harpe, P.; Cantatore, E.; van Roermund, A., ”A 10b/12b 40 kS/s SAR ADC With Data-Driven Noise Reduction Achieving up to 10.1b ENOB at 2.2 fJ/Conversion-Step,” Solid-State Circuits, IEEE Journal of , vol.PP, no.99, pp.1,8, 0

Chapter -5 Conclusion

This thesis presents a new technique called LDE-aware LCR for ultra low-power and low-frequency circuit design. Based on analyzing of low-power issues, we conclude that the leakage current, especially the STI induced junction leakage dominates the power consumption below 65nm process, especially in low frequency circuitries.

In this thesis, in chapter one, firstly, we introduce the background of requirement of low-power design. Then we analysis the major source of power dissipation in nowadays CMOS ICs. Hence, we realize the leakage, especially controlling the junction leakage is a big interest for the scaling down CMOS process.

In chapter two, the detail of LDE-aware LCR technology is proposed in for junction leakage reduction. We analyzed STI induced junction current model. Based on the analysis, we concluded and proposed LDE-aware LCR design procedure. The investigation of single transistor I-V characteristics and simulation of Inverter power consumption support our point of view.

In chapter three, LDE-aware LCR is used for an ultra-low nano-watt reference circuit design. By using LDE-aware LCR, the transistor to improve performance of sensitive power supply voltage can be saved. Moreover, we propose a novel retargeting methodology by adopting an ACM (advanced compact MOSFET) model to describe the drain current consistently in strong and weak inversion levels. In our case studies, we retarget a circuit from 1.8V/180nm process to 1.2V/90nm, 3.3V/90nm and 1.8V/65nm processes respectively. Besides, we fabricate the reference circuit in the 1.2V/90nm

process, and confirm that good measurement results are obtained with less than 12.8%/V supply voltage variation and only 1.1nW power consumption. In chapter four, for the practicality of LDE-aware LCR, we emphatic tackle the dedicated leakage reduction technique for the performance of a specific 11-bit SAR ADC. In the practical example of SAR ADC, we discuss the design procedure step by step and implement LDE-aware LCR in fundamental blocks. Moreover, we introduce the basic design specifications and methodologies for the fundamental blocks of SAR ADCs. There are a comparator energy model and a liner bootstrapped switch among them from our previous work. Some state-of-the-art low power design technologies also apply in this 11-bit SAR ADC design. The simulated and measured results convinced us that LDE-aware LCR has possibly been extended to all of low-power, low-frequency circuits. More wide verification of the feasibility by LDE-aware LCR is our future targets.

Publications

Journal

[1] Gong Chen, Qing Dong, Shigetoshi Nakatake, Zhangcai Huang, Yasuaki Inoue, “A Novel Retargeting Methodology in Computer Aided Design of Nano-watt CMOS Reference Circuit based on Advanced Compact MOSFET Model”, Journal of Computational Information Systems, To be published (Schedule: Year:2015 Volume:11 Number:5)

[2] Gong Chen, Yu Zhang, Qing Dong, Mingyu Li, Shigetoshi Nakatake, “Layout Dependent Effect-aware Leakage Current Reduction and Its Application to Low-power SAR-ADC”, IEICE Transaction on Fundamentals, To be published (Schedule: July, 2015)

[3] L i, M., Li, C., Chen, G., Zhang, Y., Dong, Q. and Nakatake, S. (2014), A New Sparse Design Framework for Broadband Power Amplifier Behavioral Modeling and Digital Predistortion. IEEJ Trans Elec Electron Eng. doi: 10.1002/tee.22002

Conference Papers

[1] Gong Chen; Bo Yang; Nakatake, S.; Zhangcai Huang; Inoue, Y., "A retargeting methodology of nano-watt CMOS reference circuit based on advanced compact MOSFET model," Circuits and Systems (ISCAS), 2012 IEEE International Symposium on , vol., no., pp.938,941, 20-23 May 2012

[2] Gong Chen; Zhang Yu; Bo Yang; Qing Dong; Nakatake, S., "A comparator energy model considering shallow trench isolation stress by geometric programming," Quality Electronic Design (ISQED), 2013 14th International Symposium on , vol., no., pp.585,590, 4-6 March 2013

[3] Gong Chen; Bo Yang; Nakatake, S., “A 9-bit 50MSps SAR ADC with Pre-charge *VCM* –based Double Input Range Algorithm,” Proceedings of the 23rd

ACM international conference on Great lakes symposium on VLSI 2013, pp. 315-316
[4] Yu Zhang; Gong Chen; Qing Dong; Mingyu Li; Nakatake, S., "Performance-driven SRAM macro design with parameterized cell considering layout-dependent effects," Very Large Scale Integration (VLSI-SoC), 2013 IFIP/IEEE 21st International Conference on , vol., no., pp.156,161, 7-9 Oct. 2013
[5] Qing Dong; Bo Yang; Gong Chen; Jing Li; Nakatake, S., "Transistor channel decomposition for structured analog layout, manufacturability and low-power applications," Quality Electronic Design (ISQED), 2012 13th International Symposium on , vol., no., pp.656,662, 19-21 March 2012