Low Power Digital Designs Operating in Subthreshold Region

SHI, Weiwei

A Thesis Submitted in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

in

Electronic Engineering

The Chinese University of Hong Kong

April 2011

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Abstracts of thesis entitled:

Low Power Digital Designs Operating in Subthreshold Region

Submitted by **SHI Weiwei**

for the degree of **Doctor of Philosophy**

in **Electronic Engineering**

at **The Chinese University of Hong Kong**

in April 2011

Subthreshold designs are required in many actual applications. Especially, the subthreshold digital systems and circuits have become more and more popular in portable devices and passive systems. In conception subthreshold digital circuits are very-low-voltage circuits, they have great reduction of power consumption but suffer from long logic delay as the driving current for logic transition and propagation is greatly reduced.

In this thesis, methodologies and examples are proposed for subthreshold digital circuit design. There is also a full study on subthreshold characteristics of devices and circuits in very-low-voltage operation. The EPC C1G2 baseband processor (BBP) for passive UHF (ultra high frequency) RFID (radio frequency identification) tag is selected as a subthreshold design example, as it is a digital design typified with instable very low supply voltage and requires ultra low power in operation. To tailor the BBP for lower operating voltage in subthreshold region, optimized structures and topologies are proposed in different hierarchical levels. In the system view, the BBP is partitioned according to the clock domain and the constraints of timing. Go down to the RTL and gate level, pipelining, parallelism, clock gating and one-hot state transition are implemented in the logic design according to the actual requirement. In this way energy awareness and power saving are achieved with enhanced robustness to operate in subthreshold region. The BBP with the proposed logic structures has

been fabricated in several deep submicron CMOS technologies. Transistor level design is the bottom level for IC designers, the proposed active control ratioed logic (ACRL) is a logic style with fast pull-up network and less capacitance, particularly suitable for the implementation of high fan-in AOI-familiar (and-or-inverter) structure. Some general ACRL cells designs, 32-bit equality comparator and, a custom PIE decoder with ACRL cells, which is the important block of BBP with critical timing, have been fabricated in 130 nm CMOS technology.

In measurement, the entire BBP design with the proposed gate-level structures exhibits high robustness in power supply and frequency variations. It can function normally at a minimum of 0.33 V power supply, which is over 100 mV below typical threshold voltage. In the test of the ACRL circuits, the ACRL cells show 30 - 70% delay reduction when compared to the standard static CMOS cells. And the ACRL custom PIE decoder works at the minimum of 0.26 V power supply, which is 40 mV lower than the minimum operating voltage archived by the PIE decoder in the BBP implemented with standard cells.

摘要

亞閾值設計有著廣泛的應用需求。尤其是亞閾值數字電路與系統，越來越受到手提式產品和無源電子系統的歡迎。在概念上，亞閾值數字電路即是超低電壓電路，它有著非常低的功耗但同時由於驅動電流劇減，邏輯信號傳遞時間變得非常長。

本論文旨在給出亞閾值數字電路設計中的可行方法及具體實例。同時亦對亞閾值晶體管與亞閾值電路設計中的特性進行全面探討。在本文中作為亞閾值設計實例的，是用於無源超高頻無線射頻識別系統標簽，並建基於 EPC C1G2 協議的基帶處理器。它的工作電壓不穩定，並且需要在超低功耗下運作，是典型的適合用於亞閾值設計的例子。為了令基帶處理器能在更低的亞閾值電壓下工作，本文在電路系統的不同的層次都提出了優化的結構和電路拓撲。在系統級設計上，整體的設計會根據不同的時鐘域和時序的緊張程度去劃分。在更下一級的門電路和 RTL 設計層面，流水線結構，並行化，時鐘閘以及單熱狀態機分別被應用在有實際需要的邏輯設計中。通過這些方法，能令在亞閾值電壓下工作的處理器更具有對能量狀態的機敏性，更節省功耗並且增強了適應性。本文所提出的處理器已經在幾個深亞微米的工藝中進行了流片。在晶體管級的電路設計是 IC 設計者的所能觸及的最底一層。本文提出一種動態控制的比例電路（ACRL）結構，它具有更快的提升網絡和更少的輸入輸出電容負載， 對於多輸入的 AOI 邏輯非常適合。在 130 納米 CMOS 工藝的投片中，部分通用的 ACRL 邏輯單元，一個 32 位的等值比較器和基帶處理器裏的重要模塊：一個 ACRL 定制的 PIE 解碼器都作為了測試樣品包含在裏面。

在投片生產後的實際測量裏，門電路級別中提出的處理器結構在電壓和頻率變化的測試環境裏展現了強健的穩定性。它的最低工作電壓是比閾值電壓少 100 多毫伏的 0.33 伏。在對 ACRL 電路的測試中，ACRL 的邏輯單元電路比標準邏輯單元電路少了 30~70%的延遲時間。同時 ACRL 定制的 PIE 解碼器能在

最低 0.26V 的電壓中正常工作，比用標準單元構造的，包含在上面提出的處理器結構裏的 PIE 解碼器低了 40 毫伏。

Contents

List of Figures

List of Tables

Acknowledgements

I have been in the ASIC/VLSI lab for over three years. From a new guy to a Ph. D candidate about to graduate, I can't make such a progress without my supervisor's help and professional guidance. Here I would like to give my deepest gratitude to Professor CHOY Chiu Sing, my respecting supervisor.

Along with Professor Choy, I would like to thank the retired Professor CHAN Cheong Fat, Professor PUN Kong Pang and Professor LEUNG Ka Nang in ASIC/VLSI lab. Your courses have widened my eye sight in IC design, your lecture classes with abundant knowledge are always charming to students. Also, I would like to thank Mr. YEUNG Wing Yee, how can I deal with EDA tools without your sincere help. Thanks to all the members in ASIC/VLSI lab, you share your knowledge with me, especially the members of RFID design group, I have learned a lot from you. And my life is colorful and enjoyable in the tennis, hiking and many other activities with you.

At last, special thanks are given to my parents and family. Happy retirement to my wonderful mother. Dozens of years' contribution to her career and her family, you are the greatest lady. Thanks to my departed grandfather, 7 years have passed, we still miss you.

CHAPTER 1 INTRODUCTION

1.1 Motivation

Dramatic increase in the demands of portable wireless applications and passive devices has stimulated the research and development on low power digital systems. In recent years, as the technology node of MOS transistors keeps scaling down to deep submicron region, with the obvious decrease in nominal power supply and the slight reduction of threshold voltage, subthreshold digital circuits have come back to the spot light of research interest. The main purpose of subthreshold design is to develop methodologies to support a longer battery life or less energy consumption. Figure 1.1 from [1] demonstrates the research trend on operating voltage as a function of year. As the whole logic core of a chip operates at a power supply falling from the nominal level to the value less than the threshold voltage of MOS transistors, the power dissipation can be greatly reduced according to the square-law dependence on supply voltage. However, the penalty of speed degradation is also obvious. Despite fabricated in deep submicron CMOS technologies, when MOS transistors in the circuit are forced to operate in subthreshold region, the drain-to-source ON current can drop much lower than 100 nA, resulting in ultra-low working frequency. Figure 1.2 sourced from [2] shows the merit and drawback of supply voltage reduction in power dissipation and operating speed respectively. Very low supply voltage also causes problems like small static noise margin, inaccurate logic cells' characterization, sensitive performance and instable functionality due to PVT (process, voltage and temperature) variation, etc.

These challenges conversely reflect the great value of the study on the alternative methods of efficient subtrheshold logic design. With the development of innovative

structures and topologies in the top-to-bottom design flow, shallow or even deep sunthreshold operation is no longer the forbidden area. It is worth taking a comprehensive look at the logic design trend.

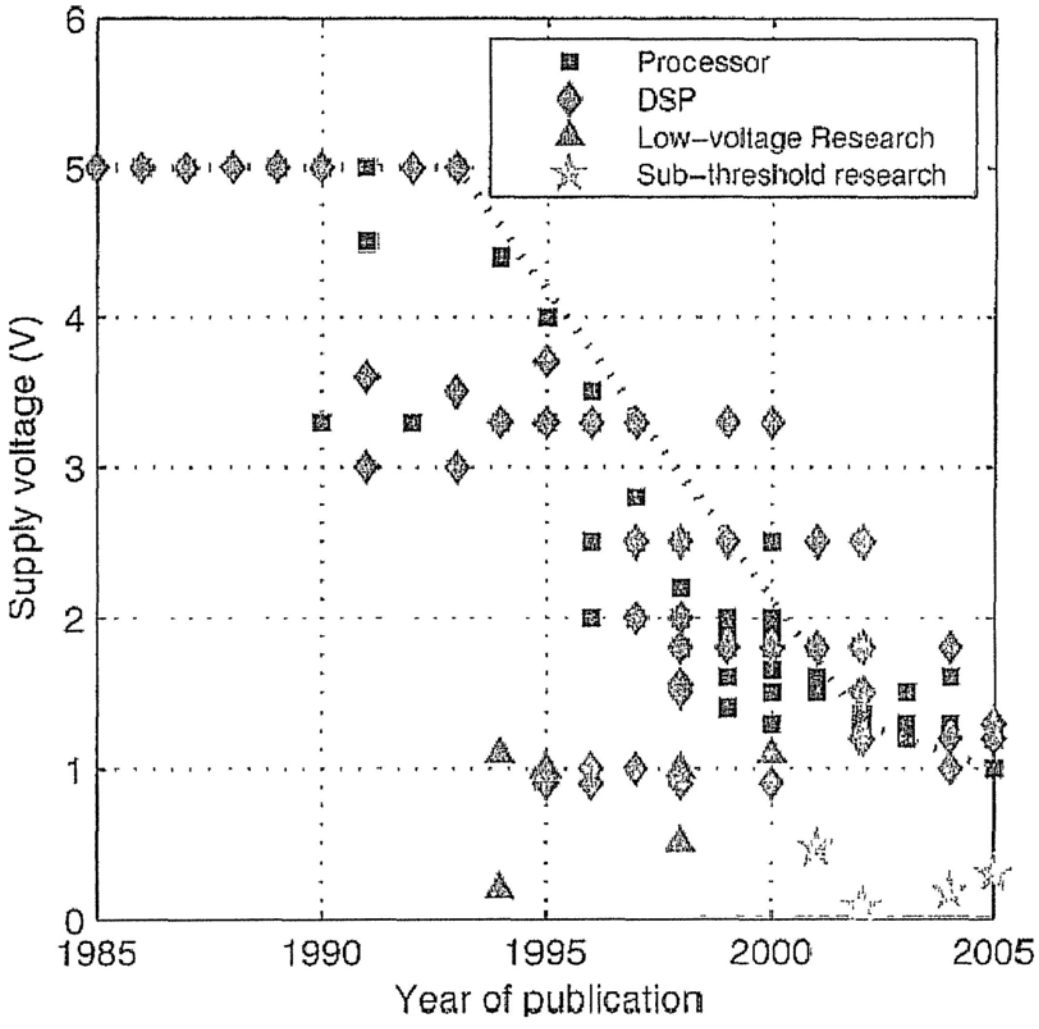


Figure 1.1 Operating voltage vs. year

1.2 Research Goals & Steps

Though the challenging may seem to be mission impossible, there are indeed several techniques, via different levels of design flow, can improve the performance and robustness of subthreshold digital circuits and systems. In another word, by the

individual or combined effort of these techniques, a target subthreshold design is achievable.

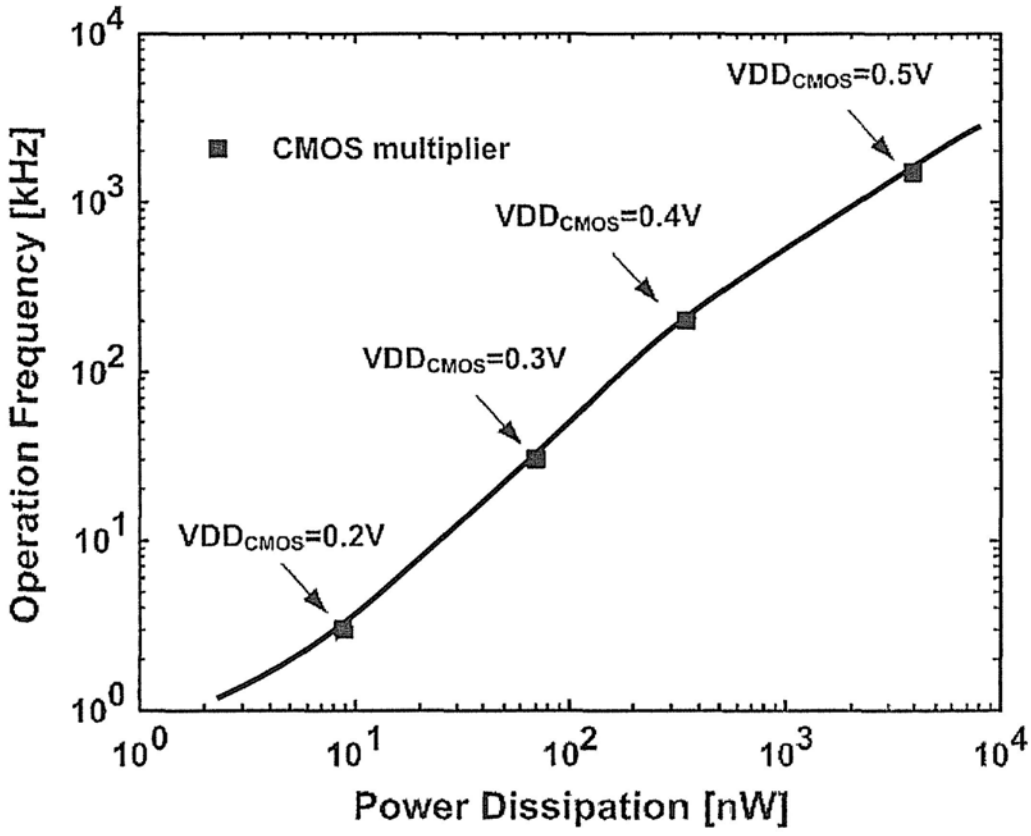


Figure 1.2 Comparison of maximum operation frequency vs. power dissipation for 8×8 bit 0.18 μm CMOS carry-save multiplier at different power supply

These techniques are mostly focused on optimizations in the aspects of system structure, datapath, digital modules, custom gate-level logic cells, configuration of MOS transistors and even the process adjustment. However a certain amount of literatures have presented research works on subthreshold logic. Early research focuses on the electrical characteristics and mechanism of the transistor in subthreshold region, as well as some ultra low frequency applications like watches. In recent years, the topic has been shifted to the robustness of subthreshold logic circuits in deep submicron CMOS technologies (180 nm, 130 nm and 90nm or below), and energy saving techniques in architecture design and gate-level circuit

design. The minimum operating voltage is also a concern in some research works. Several papers have proposed structural optimizations to reduce very long critical path delay in subthreshold operation, or have proposed novel topologies to speed up the transition propagation in the level of logic gates or compound logic circuits, which are the primary units of the whole digital system.

In this thesis, the research is carried out with several purposes to contribute to the over-all subthreshold digital IC design:

(1) At system level, based on a target design with specific requirements, efficient architectures and analysis are proposed to tailor a structure more ready for subthreshold operation. In this thesis, we select the baseband processor in passive UHF (ultra high frequency) RFID (radio frequency indentification) tags as the target example design. The chip of tag is merely powered by electromagnetic wave, and it is absolutely constrained by limited energy. Low voltage and low power are necessary.

(2) In the sub-modules of the baseband processor, studies on the critical path delay and the timing characteristics of datapath are presented to optimize the internal logic topology case by case, helping the whole circuit immune to parameters variations.

(3) According to previous research on MOS devices and logic circuits in subthreshold operation, numerical and transistor-level analysis are done. Based on the actual needs, some innovative generally-used subthreshold logic cells and BBP-required compound circuits are proposed and implemented individually, giving a full view on the performance improvement.

1.3 Thesis Organization

In Chapter 2, fundamentals and studies of subthreshold MOS design is presented. In Chapter 3, the protocol of baseband processor in UHF RFID tag is introduced in details, as well as the systematic design. Design on logic structure for low voltage and low power is discussed in Chapter 4, the whole methodology is based on the baseband processor. And as the key components of the processor, the design on PIE decoder and uplink module for wide tolerance are also proposed. Accustomed subthreshold logic cells and compound circuits are presented at length in Chapter 5. The methodology of verification and measurement, and the analysis of experimental results are in Chapter 6. Chapter 7 gives a full conclusion on this thesis.

CHAPTER 2 STUDIES OF SUBTHRESHOLD MOS DESIGN

Sub-threshold here is defined as a voltage level smaller than the threshold voltage (V_T) in a regular CMOS sub-micron technology (0.4~0.6V). Sub-threshold logic circuit is a circuitry that operates at a power supply (V_{DD}) level in sub-threshold region. Nowadays the research & development interest are focused on the deep sub-micron CMOS technology for an increasing demand of ultra low power, as the scaling down of V_T is modest to V_{DD}. As shown in Figure 2.1 [3], when the essential dimension of the technology nodes, the channel length of MOS transistor (the x axis in Figure 2.1) keeps shrinking from generation to generation, the featured nominal operating voltage of circuits follows with obvious reduction, but the threshold voltage V_T seems to be much "slower" to catch up. V_T is constrained by the physical limit of silicon process. The thickness of the gate oxide is unable to scale down as it will be too thin to maintain the reliability of this layer in stress. An over-thin gate oxide is also tend to be broken down at a slightly-high applied voltage. Therefore subthreshold operation is somehow inevitable.

The ultra-low operating V_{DD} brings all the challenges to sub-threshold design, but it is abundantly gained from energy saving. As shown in Equation (2.1), it is the equation of the dynamic power dissipation in a digital circuit or system. β is the logical transition activity factor, representing the average number of times the nodes make a logical switching in the system. C_L is the effective load capacitance: the sum-up of the voltage switching nodes related to β. f_{clk} is the working clock frequency of the system.

$$P_{dyn} = \beta \cdot {V_{DD}}^2 \cdot C_L \cdot f_{clk} \qquad (2.1)$$

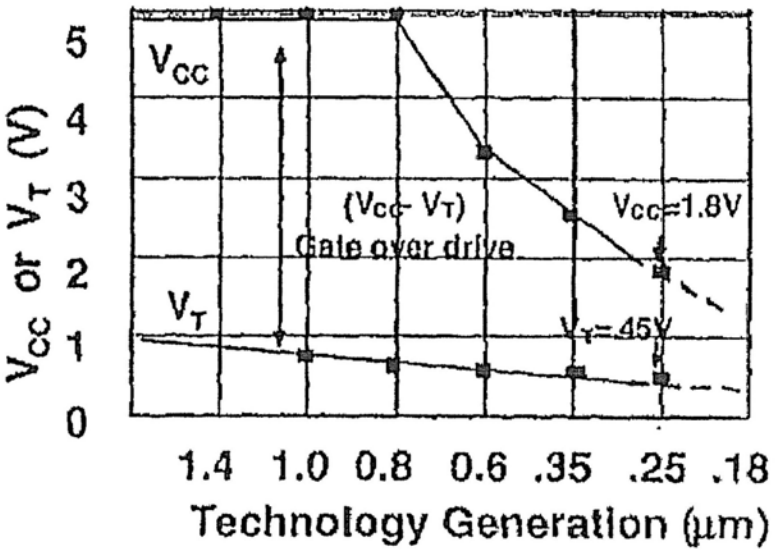


Figure 2.1 Nominal operating power supply and threshold voltage vs. dimension length of CMOS technology node

From Equation (2.1) the benefit of reducing V_{DD} is quite clear. The square dependence on V_{DD} makes an efficient decreasing of the dynamic power dissipation. However, the challenges are also hard to avoid. They bring users weak logic circuits or logic gates. The following content of this chapter will introduce the characteristics of subthreshold operation in transistor, gates and complex circuits. There are some state-of-art techniques are involved as examples. Discussions based on the target CMOS technology are also carried out in this chapter.

2.1 Transistors in Subthreshold Operation

The drain-source current of an subthreshold "ON" MOS transistor is weak, resulting weak driving ability of the gate. In nominal voltage operation, the drain-source current of MOS transistors follow the square law dependence on the gate-to-source voltage V_{GS}, if the drain-to-source voltage V_{DS} drives the transistor into saturation region. Equation (2.2) expresses the detail for NMOS transistor:

$$I_D = \frac{\mu_n C_{ox}}{2}\frac{W}{L}(V_{GS} - V_{th})^2(1 + \lambda V_{DS}) \tag{2.2}$$

where μn is the charge-carrier effective mobility, W is the gate width, L is the gate length and Cox is the gate oxide capacitance per unit area. Vth is the threshold voltage of the transistor. The additional factor involving λ, the channel-length modulation parameter, models current dependence on drain voltage due to the Early effect, or channel length modulation. The drain-source current dependence on V_{GS} of the NMOS transistor becomes linear if V_{DS} is small enough to bring the transistor into triode region:

$$I_D = \mu_n C_{ox}\frac{W}{L}\left((V_{GS} - V_{th})V_{DS} - \frac{V_{DS}^2}{2}\right) \tag{2.3}$$

The above operation region is what we can see from Figure 2.2 in [9].

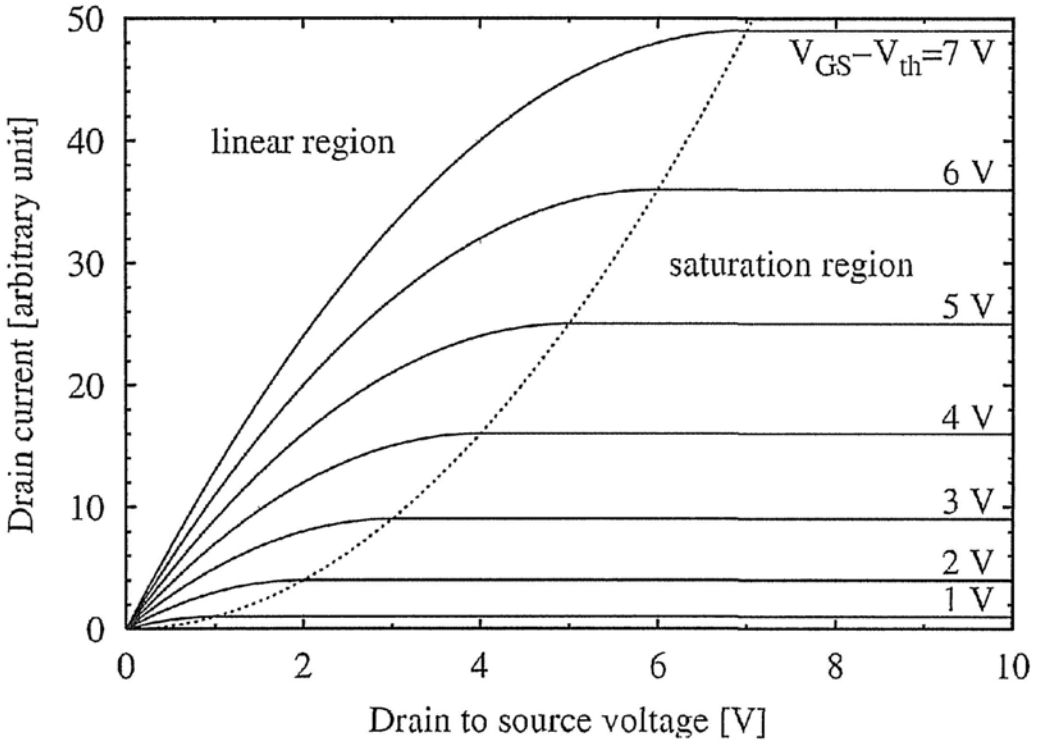


Figure 2.2 Nominal operating power supply and threshold voltage vs. dimension length of CMOS technology node

However, when the VDD is down to the threshold level, which means the swing of V_{DS} and V_{GS} are both within the range of approximately 0.4 V in deep submicron technology, the modeling for triode region is not suitable anymore. The drain-source current dependence on V_{GS} is now more sensitive, it is an exponential relationship in approximation [1][4][5][11]:

$$I_D = I_S \exp\left(\frac{V_{GS} - V_{th}}{n V_T}\right)\left(1 - \exp\left(\frac{-V_{DS}}{V_T}\right)\right) \tag{2.4}$$

where I_S = current at V_{GS} = Vth:

$$I_o = \mu_o C_{ox} \frac{W}{L}(n-1)V_{th}^2 \tag{2.5}$$

and the slope factor n is given by:

$$n = 1 + C_D / C_{OX} \tag{2.6}$$

with C_D is capacitance of the depletion layer and C_{OX} is capacitance of the oxide layer. V_T here is the thermal voltage that is defined by:

$$VT = kT / q \tag{2.7}$$

where q is the magnitude of the electrical charge on the electron with a value $1.602\,176\,487 \times 10^{-19}$ C. In electronvolts, the Boltzmann constant k is $8.617\,343 \times 10^{-5}$ eV/K. Figure 2.3 shows the exponential dependence in a deep submicron technology [1].

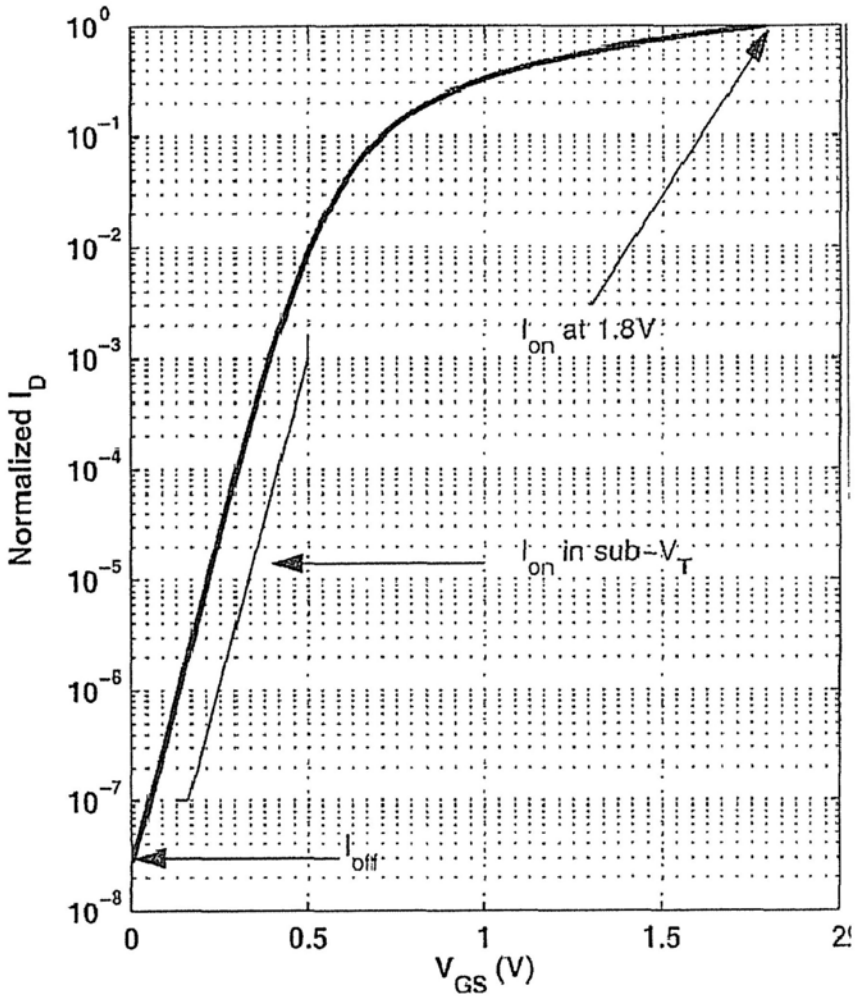


Figure 2.3 Drain-source current vs. the gate-source voltage ($V_{DD} = V_{GS}$)

From Figure 2.3 we can recognize that the drain-source current decrease rapidly as the VGS is reduced towards zero. For subthreshold research, there is an important parameter called *subthreshold slope S*. *S* gives the inverse of the slope of I_D versus V_{GS} in millivolts per decade of change in I_D. It can be derived from Equation (2.4), (2.6) and (2.7):

$$S = n V_T \ln 10 \tag{2.8}$$

This parameter tells the ability of the gate-source voltage to change the drain-source current in subthreshold region. For a smaller subthreshold slope S, the device obtains

a larger subthreshold I_{ON}/I_{OFF} ratio. As the subthreshold I_{ON} is comparatively larger and I_{OFF} is smaller, the robustness of logic gates is improved: the Voltage Transfer Curve (VTC) can be steep and the leakage current can be very low. Together with the threshold voltage Vth, they can define the specific amount of I_{ON} and I_{OFF} in subthreshold operation.

Besides the merit of ultra-low power, subthreshold operation also suppresses the effects of short-channel and DIBL (drain induced barrier lowering), as the subthreshold drain-source current are mainly influenced by V_{GS}. In the research field or industries, there are several methods to better these merits of the transistor in subthreshold operation.

2.1.1 Process Optimization

In the process level, a modified doping technique in the active area of MOSFETs can change the subthreshold slope and the current characteristics. It is an established fact that for scaled super-threshold transistors it is essential to have halo andretrograde doping to suppress the short channel effects. The main functions of halo doping and retrograde wells are to reduce drain-induced barrier lowering (DIBL), prevent body punchthrough and control the threshold voltage of the device independent of its subthreshold slope. However, in subthreshold operation, it is worthwhile to note that the overall supply bias is small (in the order of 150~ 300 mV). DIBL and short-channel effect are very slight. Hence, it can be qualitatively argued that the halo and retrograde doping are not essential for subthreshold MOS transistor design [6].

The absence of the halo and retrograde doping has the following implications:

- *A simplified process technology in terms of process steps and cost.6*

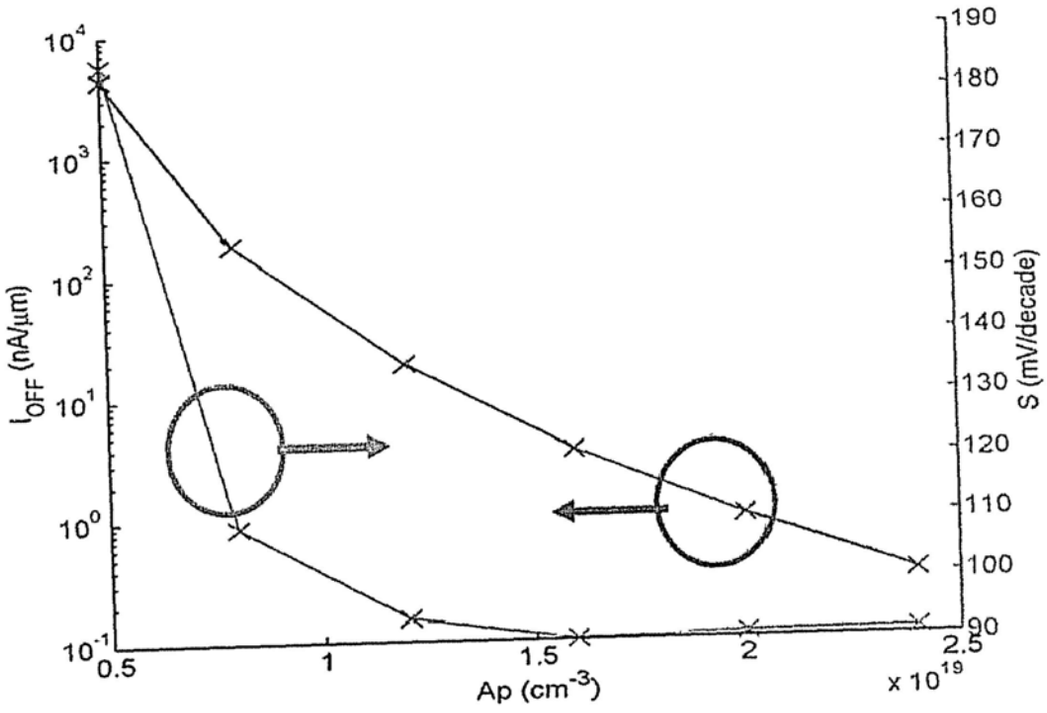
- *A significant reduction of the junction capacitances resulting in faster operation and lower power.*

And if we optimize the doping with a process-simplified high-to-low profile, It will have a low doping level in the bulk of the device to:

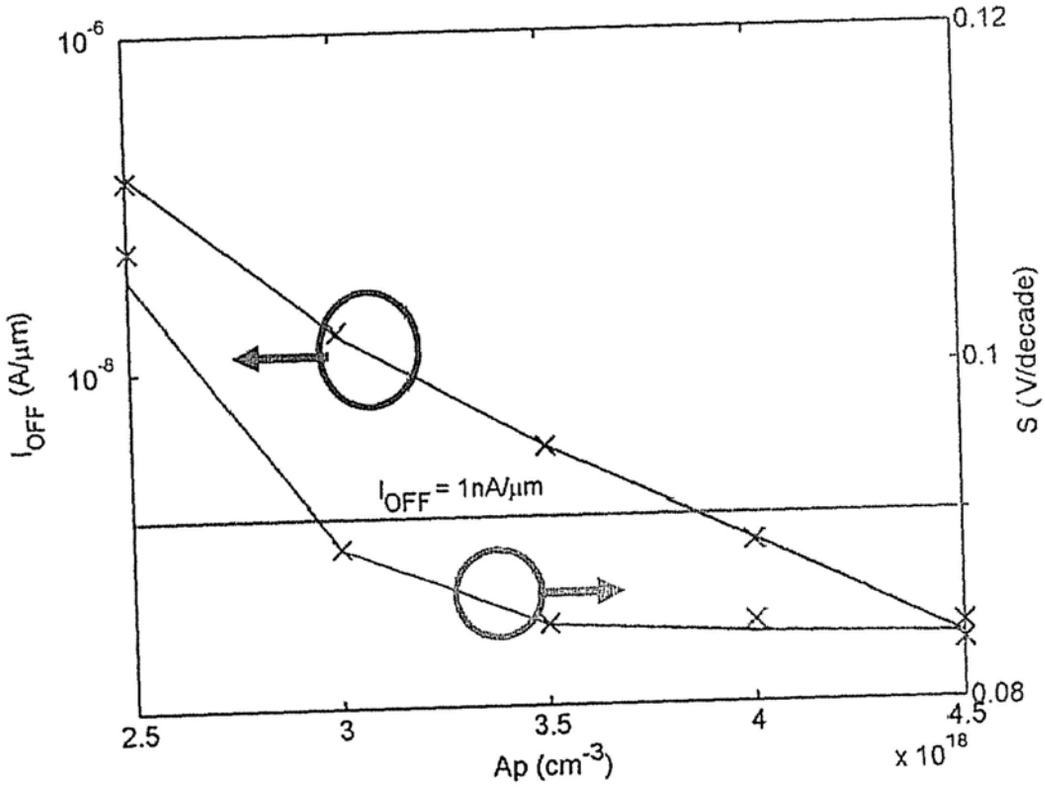
- *Reduce the capacitance of the bottom junction.*

- *Reduce substrate noise effects and parasitic latch-up problems.*

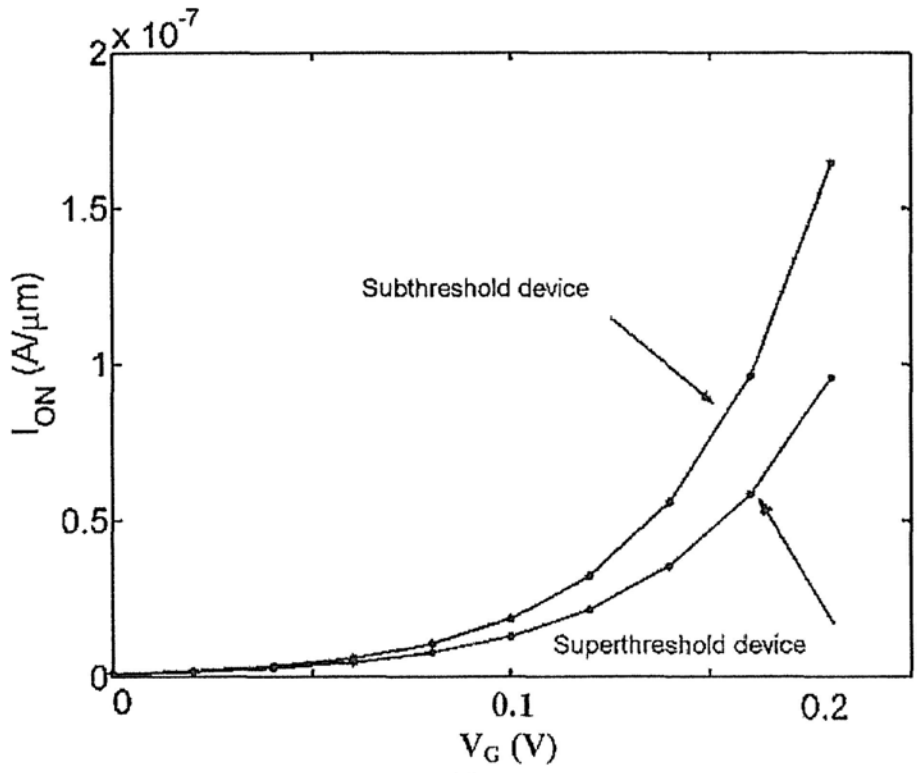
Figure 2.4(a) and (b) from [6] demonstrate the improvement of the optimized device in 50 nm technology. With the comparison we can see the high-to-low profile has a smaller subthreshold slope and smaller I_{OFF} current. Additionally, at the same level of I_{OFF} current, the optimized subthreshold device in [6] offers higher I_{ON}. Figure 2.4(c) demonstrates how these combined efforts take the effect on I_{ON}.


10^4
10^3
10^2
10^1
10^0
10^{-1}
I_{OFF} (nA/μm)
190
180
170
160
150
140
130
120
110
100
90
S (mV/decade)
0.5
1
1.5
2
2.5
Ap (cm^{-3})
x 10^{19}

(a)


10^{-6}
10^{-8}
I_{OFF} (A/μm)
0.12
0.1
0.08
S (V/decade)
I_{OFF} = 1nA/μm
2.5
3
3.5
4
4.5
Ap (cm^{-3})
x 10^{18}

(b)



(c)

Figure 2.4 (a) Variation of I_{OFF} and S with the peak of the halo doping, Ap. (b) I_{OFF} and S for the high-to-low doping profile for varying Ap. A 200 mV V_{DD} has been chosen here. (c) I_{DS} - V_{GS} for the super-threshold and the subthreshold device for iso-IOFF (1 nA/μm).

2.1.2 Body Bias in Subthreshold Operation

To change the performance or I-V characteristics in a MOS transistor, the most efficient way is the adjustment of the transistor's threshold voltage. It can be implemented by modifying the doping concentration of the substrate of the device, implementing impurity near the silicon surface under the gate, or by increasing / decreasing the thickness of the gate oxide. These are the effective ways to change Vth. Other alternative is the sizing of transistors. It is the main practical method to change the current capability, and it also has a small impact on threshold voltage. This will be discussed later with logic gate design in subthreshold region.

If the process of the technology is fixed, gate oxide thinning and doping techniques are not available. Another commonly used techniques is body bias, it also has an obvious adjustment to device performance in subthreshold operation. Body bias was originally used in analog circuit design, when the power supply is reduced to low voltage level, body bias effect is attracting logic designers' attention, especially when their target design is in deep submicron technology. Figure 2.5 is the cross-section of a NMOS transistor. Body bias uses the voltage drop applied between the source terminal and the bulk in the MOS transistor, to modify the conductive channel's sensitivity to the gate-source voltage drop. Equation (2.9) gives the function of V_{TO}, the threshold voltage without the influence of substrate bias, but the impact of gate oxide thickness and doping concentration [7][23]:

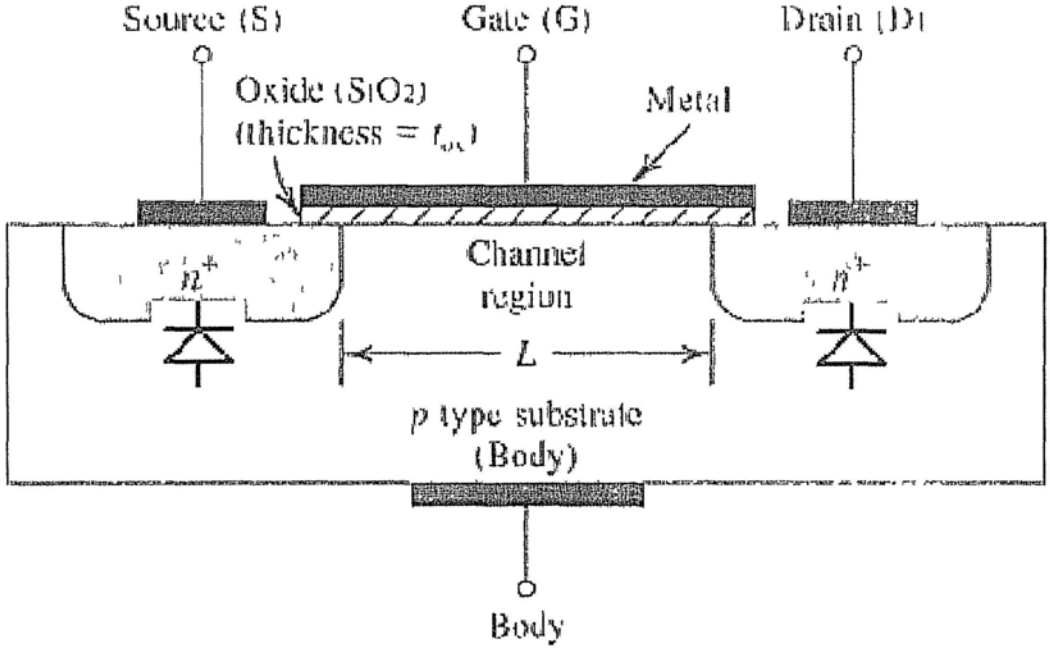


Figure 2.5 Cross-section of NMOS device

$$V_{T0} = -\phi_{ms} - 2\phi_F + \frac{Q_{b0}^{'} - Q_{ss}^{'}}{C_{ox}^{'}} = V_{FB} - 2\phi_F + \frac{Q_{b0}^{'}}{C_{ox}^{'}} \quad (2.9)$$

Equation (2.10) gives the expression that how body bias can affect the threshold voltage Vth, based on the original V_{T0}. Here NMOS transistor is taking as an example [1][7][23]:

$$V_{TN} = V_{TO} + \gamma(\sqrt{V_{SB} + 2\phi_F} - \sqrt{2\phi_F}) \tag{2.10}$$

where V_{TN} is the threshold voltage when substrate bias is present, V_{SB} is the source-to-body substrate bias, $2\varphi F$ is the surface potential, and V_{T0} is threshold voltage for zero substrate bias, the body effect parameter γ, is given in Equation (2.11) as below:

$$\gamma = (t_{ox}/\epsilon_{ox})\sqrt{2q\epsilon_{si}N_A} \tag{2.11}$$

where t_{ox} is oxide thickness, ε_{ox} is oxide permitivity, ε_{si} is the permitivity of silicon, N_A is a doping concentration, q is the charge of an electron. Figure 2.6 shows the body bias effect on the threshold voltage in a 90 nm CMOS technology, with a nominal Vth of about 0.4 V [8].

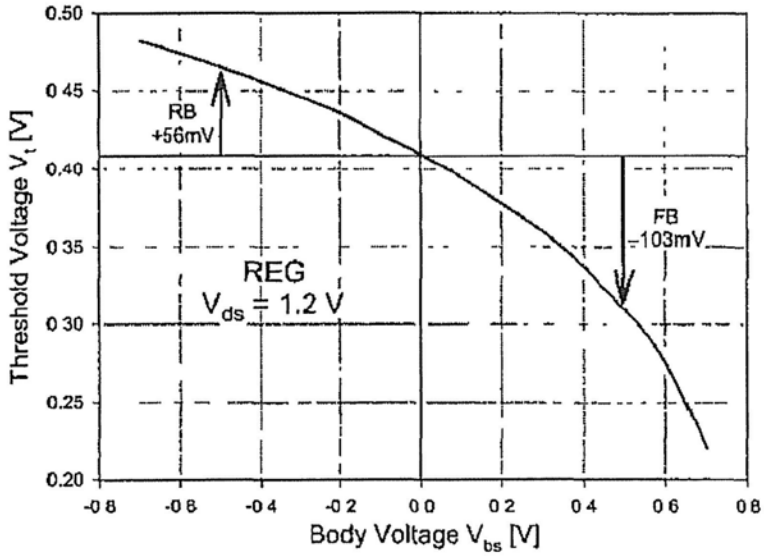
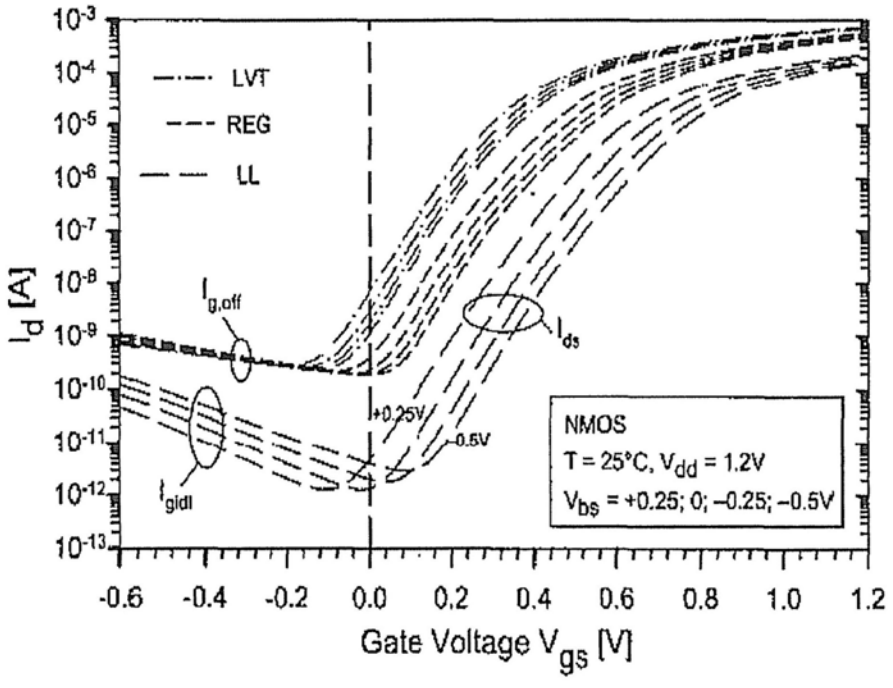


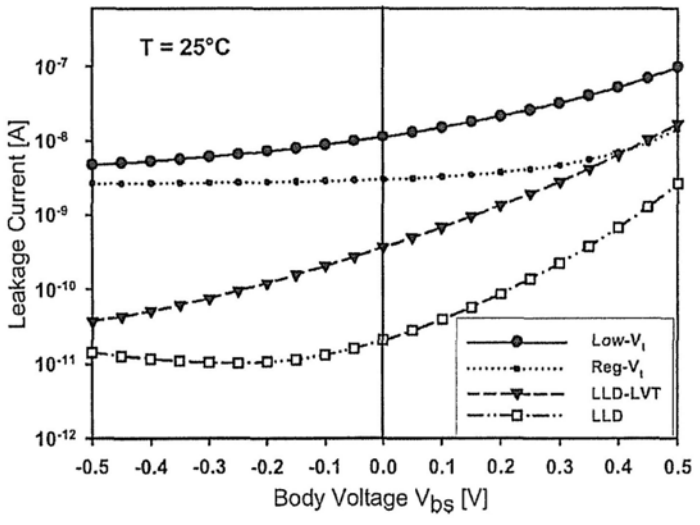
Figure 2.6 Body bias effect on threshold voltage

As mentioned before, the drain-source current in subthreshold operation has an exponential dependence on the gate-source voltage Vgs. Therefore body bias can

obviously increase the drain-source current no matter the device is turn ON or shut OFF with only leakage current. Figure 2.7(a) and (b) shows the current variation with the change of body bias in different respects of view [8].



(a)



(b)

Figure 2.7 (a) ON current vs. Vgs with different body bias and device. (b) OFF current vs. body bias with different device

It is clear that body bias takes effect on devices with different threshold voltage.(Low-Vt stands for low threshold voltage, Reg-Vt is regular threshold voltage, and LLD means low-leakage device). And it is supposed to be a popular technique in building up a subthreshold circuit. However, as the subthreshold operation is indeed a leakage current operation with variable magnitude, the side effects of body bias can be amplified to ruin the over-all performance. The first one is the increased leakage current in OFF state, resulting in additional static power consumption.

The second one is the most possible factor that degrades body bias in subthreshold: the increased junction capacitance of the bulk-to-active-area. Recall Figure 2.5, we can recognize that the forward biased V_{bs} is indeed a forward biased diode at the interface of source area and the P-type substrate. In this condition the junction capacitance is larger than the zero-bias one, due to the reduced effective distance between the two charge layers in the depletion region of both sides. This reduction is caused by the forward-bias as Equation (2.11) and (2.12), (2.13) describes the mechanism:

$$C = \frac{A\,\varepsilon}{W} \tag{2.11}$$

A is parallel plates of area, W is the width of plate separation, and the permittivity is ε. If we go further on the W, we will find [10][12]:

$$W(V) = (V_o - V)^{1/2}\left(\frac{2\,\varepsilon_r\,\varepsilon_o}{q}\right)^{1/2}\left(\frac{N_a + N_d}{N_a N_d}\right)^{1/2} \tag{2.12}$$

where Vo is the built-in potential of the device (a constant). Na and Nd are the doping concentration of each side. V is the applied voltage on the device. The junction capacitance then becomes as below:

$$C_j(V) = A\left(\frac{q\varepsilon_r\varepsilon_o}{2}\frac{N_a N_d}{N_a + N_d}\right)^{1/2}\frac{1}{|V_o - V|^{1/2}} \quad (2.13)$$

From Equation (2.14), the influence of the bias voltage V is inevitable. When forward bias is applied, (Vo-V) increases, and Cj reduces. Figure 2.8 indicates the effect in curve plotting. It shows the capacitance increasing of the two components of pn junction.

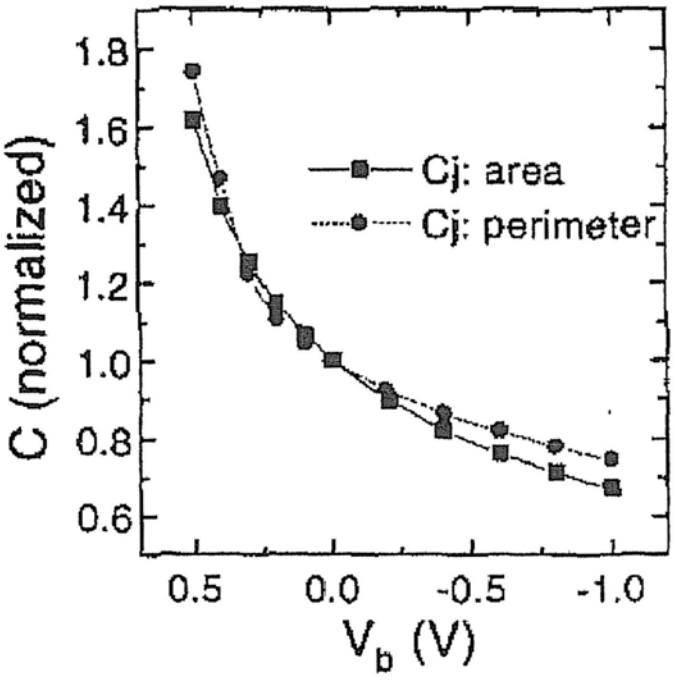
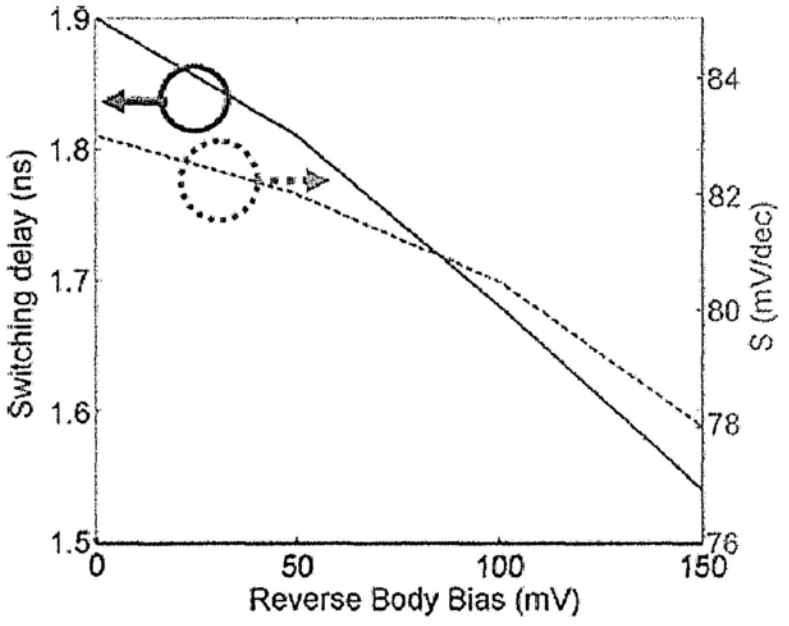


Figure 2.8 Junction capacitance (divided into two components) vs. the body bias voltage (forward biasing is positive)

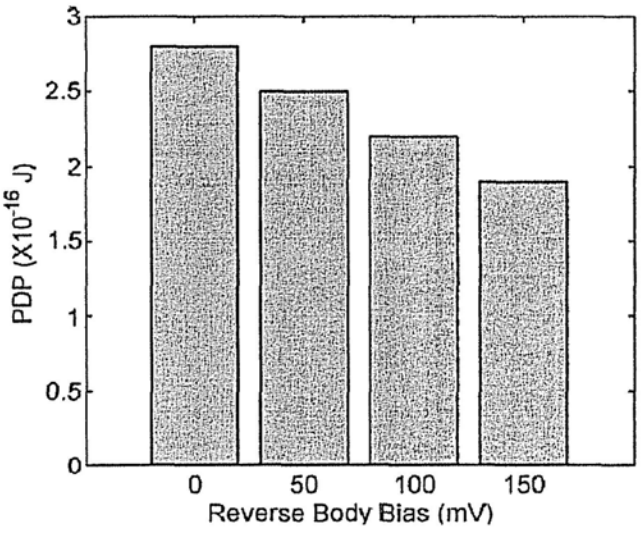
Therefore, though the ON current is increased, the additional junction capacitance is there. Especially in subthreshold operation, the ON current is in fact a leakage current so the body-bias-induced current increases from a very low amount. As the propagation delay of a logic gate is dependent on both of the driving current of ON device and the load capacitance, which includes the junction capacitance of the device, it has to analysis case by case to determine the efficiency of forward body bias.

The third one of the negative effects are the short channel effect (SCE). Forward biasing reduces the threshold voltage and increases the instability in the control of channel.

In conclusion, whether body bias can improve the circuit performance depends on the electrical characteristics of the process technology. If the increment of the device's current, which is induced by forward body bias, can fight over the increment of junction capacitance caused by the same bias, then we can say forward body bias is worth of applying in this technology. Like the design in [8] using a 90 nm technology. If the current increasing can't overcome the increased capacitance, body bias may degrade the circuit design and waste chip area. And conversely, as shown in Figure 2.8 and concluded from Equation (2.13), even though reverse body bias reduces the ON current, it also reduces the junction capacitance. If the capacitance reduction has more impact than the current degradation, a reverse body bias logic gate can even have a better performance. Figure 2.9(a) from [13] gives such an example of inverter, which is fabricated in 50 nm CMOS technology. We can see that both of the logic delay and subthreshold slope S are reduced as the bulk-to-source voltage is decreased from zero. Additionally, reverse body bias can reduce the OFF current and consequently reduce the static power. In a proper process, reverse bias increment can decrease the power delay product (PDP) from both of the components. Figure 2.9(b) demonstrates the trend.



(a)



(b)

Figure 2.9 (a) Reverse body bias effect on logic delay and subthreshold slope. (b) Improvement of power delay product by reverse body bias

2.2 Subthreshold Logic Gates and Circuits

Subthreshold logic gates and circuits can be implemented in different topology styles. In fact, many logic styles can function at a subthreshold power supply. But it requires more robustness, efficiency and simplified scheme of device stacking to ensure the reliability of the design, especially the transistors are weak of current driving in very low voltage condition.

Main stream of subthreshold MOS logic research focus on the improvement of the conventional logic styles. Most attention are on the static CMOS logic gates, which are commonly used in commercial market and with well developed topology. Because it is a full swing signal transition, static CMOS style has a better robustness, and the static power is very low as the leakage current is suppressed by at least one shut-OFF transistor.

Deriving from static CMOS style, pseudo NMOS logic is a fast-response style. Compared to static CMOS, the pull-up network is replaced by a single pull-up active PMOS load, and the pull-down NMOS network remains. Due to the simplified input capacitance and pull-up path, pseudo NMOS logic is generally considered faster than static CMOS, however, as the pull-up path is always active, the static power is dramatically increased when there is at least a path from output to the ground is ON. And careful sizing of transistors is required in this logic style [47], otherwise the functionality and the speed of the gate is out of guarantee.

Another alternative is a mixed style which speeds up the logical transition with the combined effort of skewed static CMOS and dynamic CMOS logic. Skewed logic is basically static CMOS logic but the sizing is with a tendency for faster low-to-high or high-to-low transition [57]. Inserted by the precharge-to-high or precharge-to-low dynamic logic gates, they form the datapaths of the whole circuit, giving a fast design

in subthreshold operation. The penalties are mainly the cost of chip area and routing complex.

Differential structure of circuit is generally used in analog field to amplified small magnitude input signal. In the past it was also implemented in logic circuits, like differential bipolar logic or MOS logic. The latter one is general called current mode logic (CML), the input-controlled MOS pull-down or pull-up network appear in a pair, and they are complementary to each other. So the differential structure is in fact a voltage decision topology with biased tail current and dual complementary input networks. This logic style is featured by controllable voltage swing, less sensitive of PVT variation and faster propagation with large tail current. But the complexity is high, as well as the static power consumption. The design of Bias circuit and the mismatch consideration also add the work load. Similar logic style named cascade voltage switch logic (CVSL) is also a candidate for subthreshold usage, it has no tail current bias and the pull-up load is cross controlled by the dual complementary input networks.

2.2.1 Static CMOS and SRAM

In static CMOS design, the sizing of the gate circuit is the most effective way to adapt the performance and robustness. Figure 2.10 shows how the ratios of the width of PMOS and NMOS transistor in an inverter of 350 nm technology can affect the delay of this gate [14]. Besides the speed of circuits, proper sizing can also help the whole system design achieve minimum voltage operation. Figure 2.11 gives the simulation example on defining the minimum supply voltage for an inverter in 180 nm CMOS technology [15][16]. It also defines the ratio sizing range of the width of PMOS and NMOS transistor at a different supply voltage. Additionally, Figure 2.12

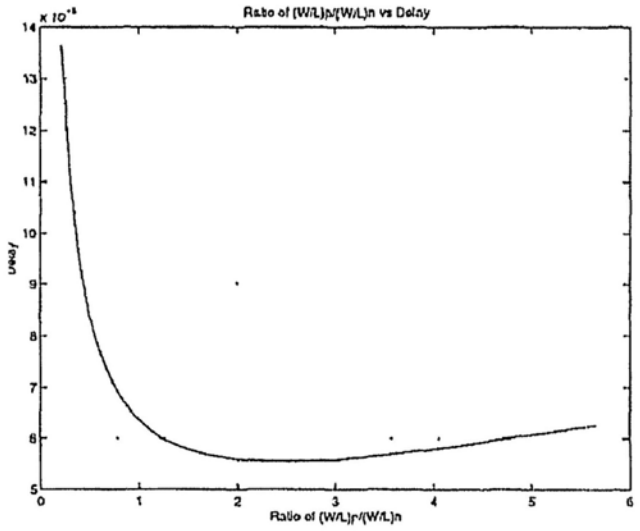


Figure 2.10 Inverter delay vs. ratio of PMOS/NMOS width

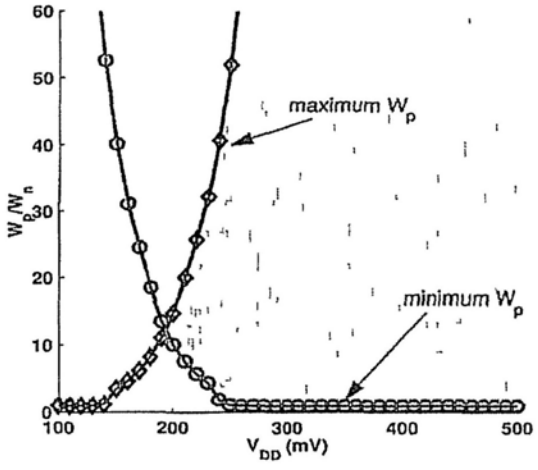


Figure 2.11 Minimum and Maximum sizing ratio of the width of PMOS and NMOS transistor vs. operating VDD in the inverter

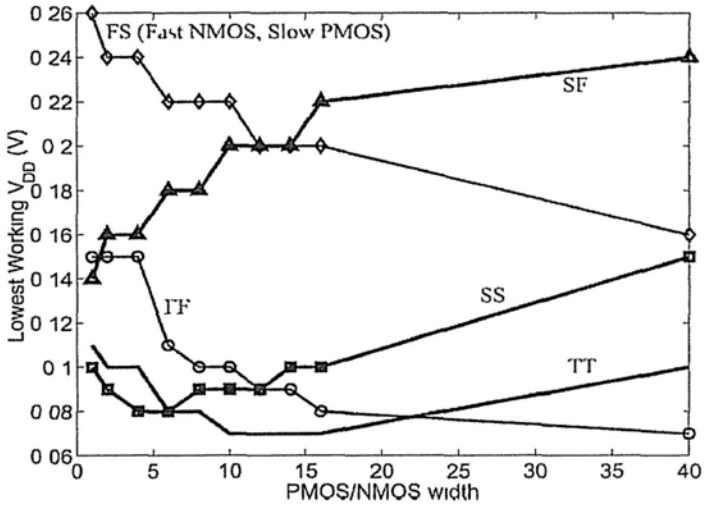


Figure 2.12 Minimum working VDD versus width ration in different process corner

pluses the process corners in consideration to provide a full view of the minimum operating voltage versus PMOS/NMOS sizing [15][16]. It is a useful method that helps the designer know the boundary and select appropriate transistors to meet the specification of the whole subthreshold design.

Figure 2.13 is a general D flip flop (DFF) structure in a standard logic cell library in 180 nm technology [15]. In realistic, the DFF can function conditionally at 300 mV VDD if the process of fabrication is within typical-typical case. However, if the fabrication has derivation in process, it is likely not workable until VDD rise up to 400 mV. The worst condition occur when the process corner is FS (fast NMOS, slow PMOS), and clock signal CK =0. The combined, strong off current (dashed lines with arrows) in the nMOS devices in I6 and I3 (larger sized) overcomes the weakened, narrow current of PMOS device in I6. At nominal voltage level, OFF leakage is nothing, but subthreshold region, the functionality is not secured. By launching the similar simulations like Figure 2.11 and 2.12, we can see that a larger size of PMOS in I6 can solve the current fighting problem at low voltage supply. After size adjustment, the DFF can work normally at 300 mV for ever process corner.

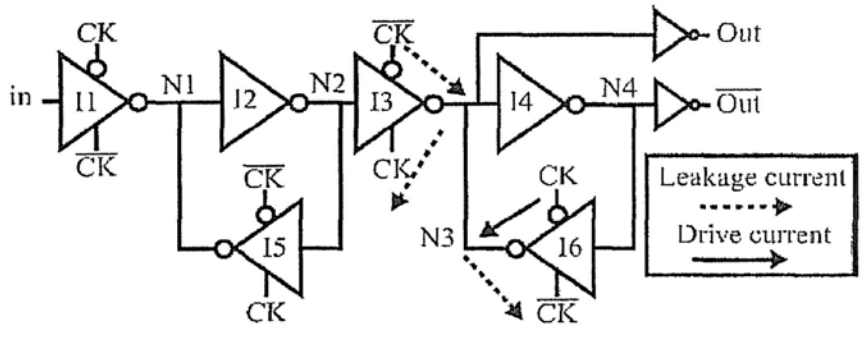


Figure 2.13 Worst case leakage condition in static CMOS DFF

Nodal OFF current reduction is another way to boost up logic gates robustness in low voltage. As the OFF branch connect to a node may contribute unwanted OFF current that destroy the supposed voltage level, OFF current reduction tries to reduce

unfriendly branch in worst-case input pattern or add ON and OFF branch current to support the target voltage level at output node. Figure 2.14(a) is a bad design in 180 nm technology of 2-input XOR gate for subthreshold operation [17]. Fatal error happens when input A =0, B=1, and the output Z is supposed to be logical 1. However, there is only one ON branch that is feeding high voltage level at the output node, and three OFF branch is fighting with it to pull Z down. At 100 mV VDD, the XOR gate fails to work. But if we modify the topology a bit, the result is totally different. Figure 2.14(b) gives the right design for very low voltage. Additional pull-up branch helps maintaining the output level even the VDD is 100 mV [17].

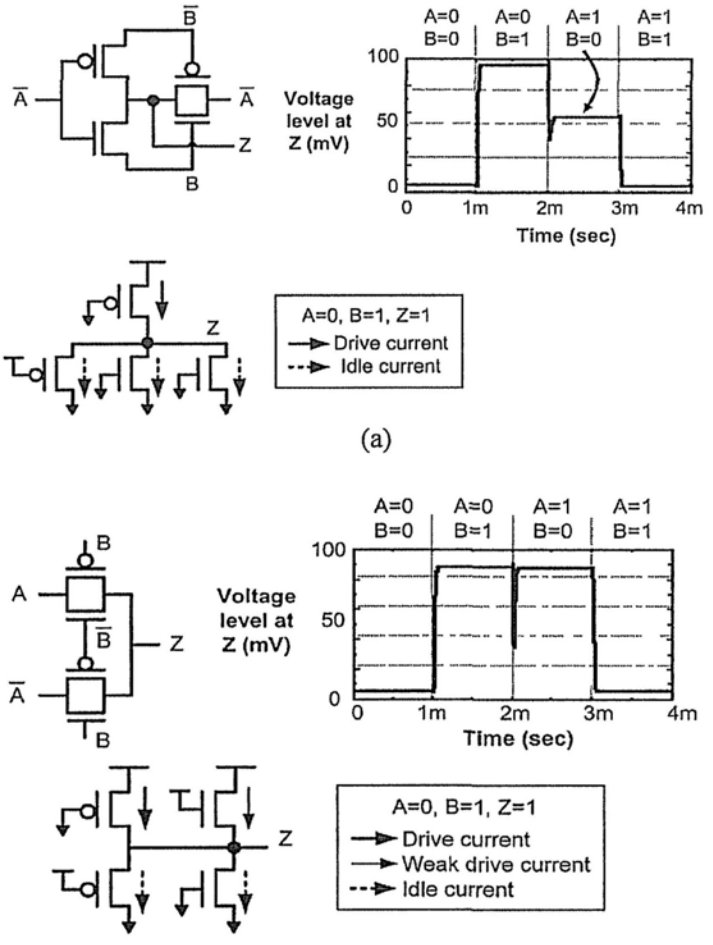


Figure 2.14 (a) bad design of XOR gate with subthreshold leakage hazard. (b) topology optimized XOR gate reduces the leakage impact

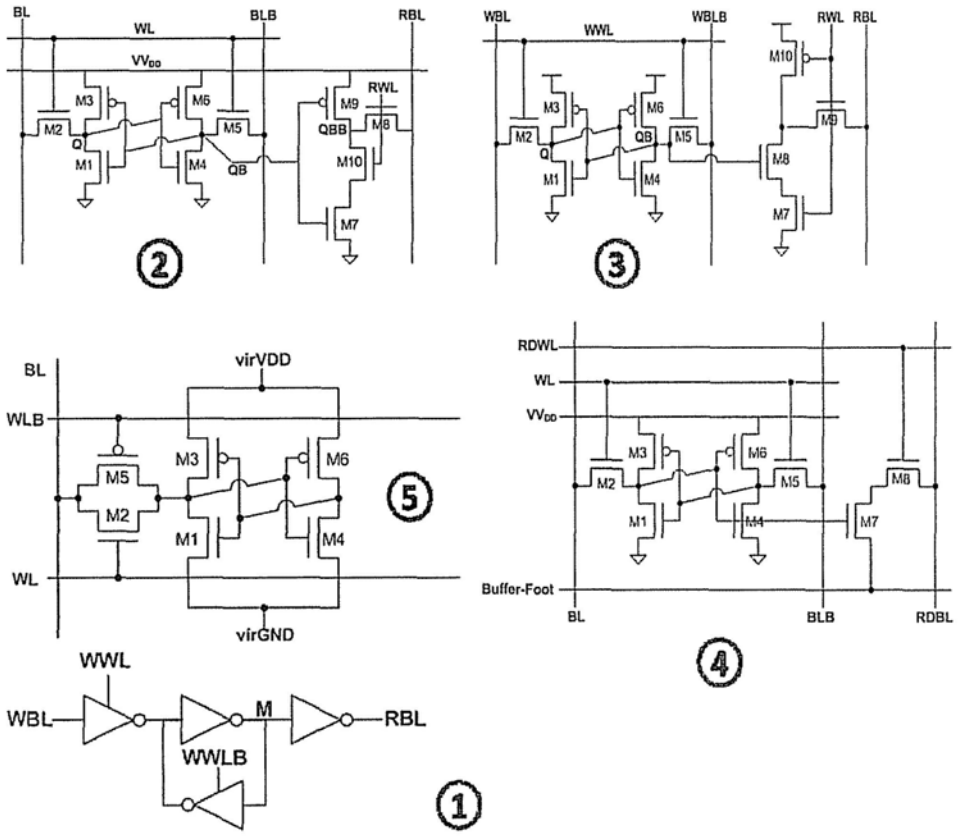


Figure 2.15 Published works of SRAM cell design with sunthreshold leakage reduction

SRAM cell is not belonged to static CMOS gates, but it is frequently used in digital systems and sometimes occupies the most chip area. The subthreshold research on SRAM cells is reported in many published papers. The general methodology is similar to the static CMOS one: to optimize the topology that reduces hazard OFF leakage current and strengthen the ON or OFF current friendly to the target output. Figure 2.15 lists the published works in recent years, and Table 2.1 gives the comparison of them [18]. For the consideration of comparatively weak driving capability of the selected single cell in subthreshold weak operation, most of the designs focus on the read leakage reduction. Like the No. 4 design in Figure 2.15, only two more NMOS transistors are added to the cell, but the hazarding leakage is isolated from the storage core of the SRAM cell. Detail routing of this cell is shown

in Figure 2.16 [19]. RDBL is set to high in read operation, the Buffer-foot is low at this time. If M7 is ON, sense amplifier will detect the fighting at RDBL, otherwise, RDBL will keep as high. In this way, the two cross-controlled core inverters are protected from status change that can be caused by the OFF leaking branches outside, especially when the number of them is over 100.

Table 2.1 Comparison of subthreshold SRAM cells

	[1]	[2]	[3]	[4]	[5]
Supply Voltage	180 mV	400 mV	200 mV	350 mV	210 mV
Process	180 nm	65 nm	130 nm	65 nm	130 nm
Reported Frequency	Not Reported	475 kHz	120 kHz	25 kHz	21.5 kHz
Transistor Count	12	10	10	8	6
Max # Rows	128	256	1024	256	16
Read SNM	54.8 mV	139.5 mV	59.5 mV	118.4 mV	44.1 mV
Read SNM/SNM6T	1.6x	2.4x	2.4x	2.5x	1.4x
Write SNM	N/A	209.3 mV	63.8 mV	153.7 mV	77.1 mV
Write SNM/SNM6T	N/A	3.9x	6.3x	3.8x	Infinite

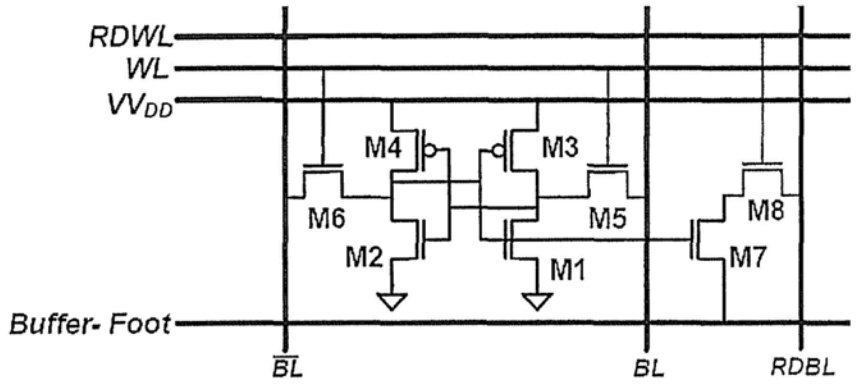


Figure 2.16 8T SRAM bit cell for subthreshold read operation

2.2.2 Pseudo NMOS

Conventional pseudo NMOS logic design is a matured technique. But the structure is lack of robustness in subthreshold operation. The sizing of PMOS and NMOS ratio

must be careful and the static power is hard to decrease. An innovative pseudo NMOS design was proposed in recent year that is combined with body bias technique for dynamic control.

As shown in Figure 2.17, the NMOS pull-down network is different from the conventional one. Each NMOS in the network shares the same substrate no more, each one's substrate is now directly connected to the input which also controls the NMOS gate. Threshold voltage control automation is embodied in the signal-decided body bias. When the NMOS is ON, Vbs rises up, Vth is reduced, so the subthreshold ON current will be higher. Hence the self control NMOS is called dynamic threshold device, such pseudo NMOS style is named DT-pseudo-NMOS. Figure 2.18 shows the power-delay-product comparison of three types of pseudo NMOS 3-input NAND gate, fabricated in 350 nm technology [20].

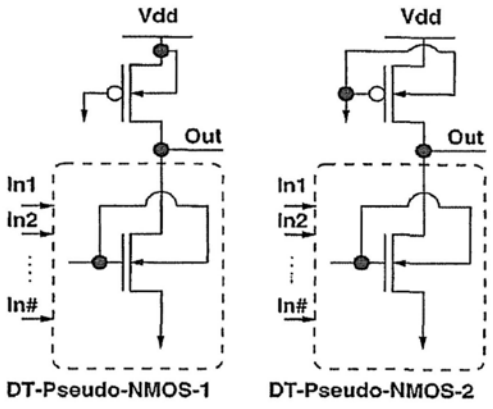


Figure 2.17 Innovative dynamic threshold pseudo NMOS logic, with input-controlled separated body bias

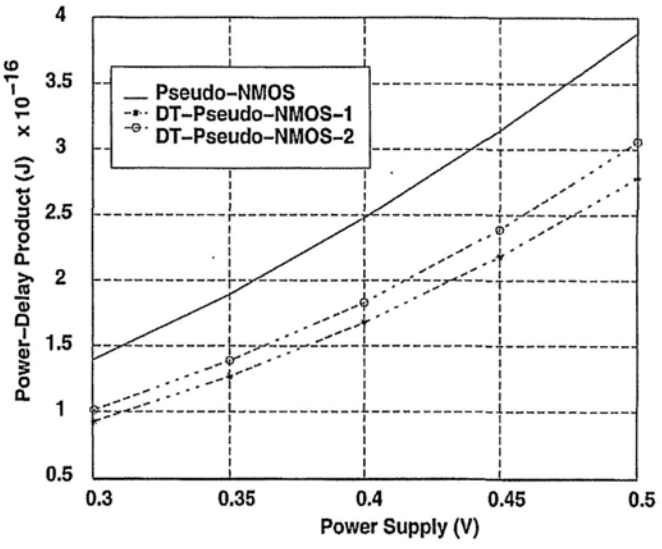


Figure 2.18 Power-delay product comparison of pseudo NMOS styles vs. power supply

From Figure 2.18 the DT-pseudo-NMOS seems to have a much better performance than conventional one. However, in lower VDD and new technologies with shorter channel length, the increased junction capacitance may eliminate the benefit. So whether this style is suitable for deep submicron process is a dependent question. Possibly there are several logic functions which are fit to be implemented, but the cost of complexity is unavoidable.

2.2.3 Current Mode Logic

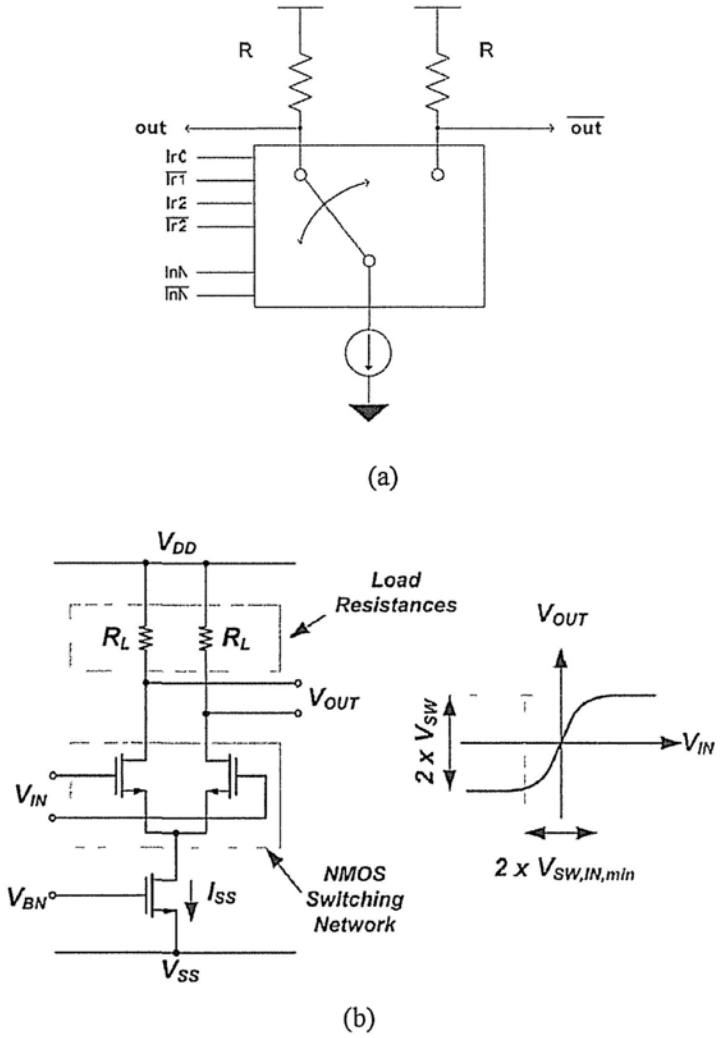


Figure 2.19 (a) Basic structure of MCML; (b) Example of a basic MCML inverter

Current mode logic has a long history in digital logic gate design. first application is the bipolar emitter coupled logic (ECL), when technology node is shifted to MOS process, MOS current mode logic (MCML) appeared. Figure 2.19(a) shows the basic structure of MCML. There are two NMOS pull-down networks which are complementary to each other. When one network in ON, the output node at this part will sink to the logical low level, and the pull-up resistor for the other network will charge its output node at logical high, giving a differential outputs. Figure 2.19(b) is the example of a basic MCML inverter [21][23].

New research on this topology has elaborated the advantages of MCML like small voltage swing, fast transition as the voltage gain is high and capacitance is smaller, and the robustness to PVT variation as the bias current is controlled by band gap reference which is self-adjustable to satisfy different conditions. Figure 2.20 gives the detail example of an inverter of the innovative design, called subthreshold source-coupled logic (STSCL) [21].

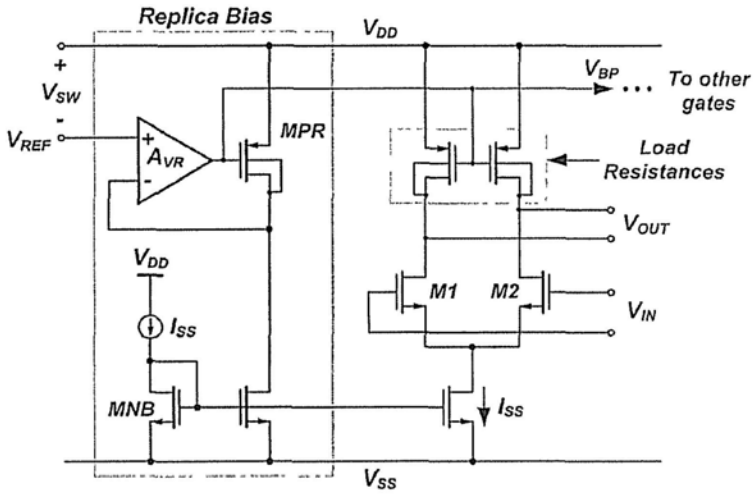


Figure 2.20 STSCL inverter with bulk-drain-connected PMOS load and bias

The pull-up resistors in the basic topology are replaced by the active PMOS resistive load. And different from the conventional PMOS load, this drain-bulk-connected one offers a better I-V character as an ohm resistor in subthreshold region. Figure 2.21 gives the comparison of this two configurations of PMOS [21][23]. Obviously a better voltage gain is obtained when drain-bulk-connected PMOS is

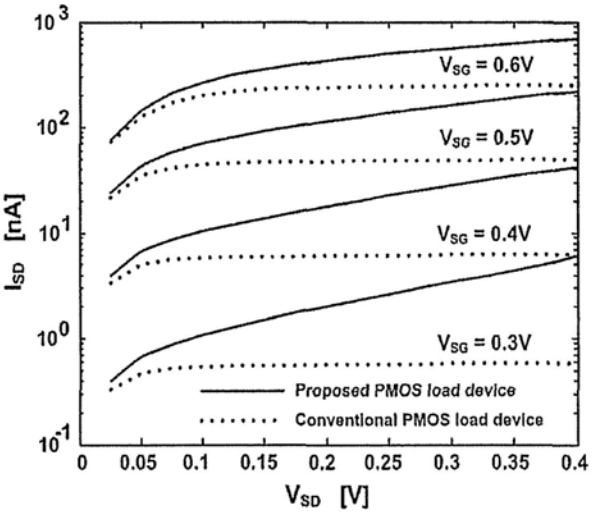


Figure 2.21 STSCL inverter with bulk-drain-connected PMOS load and bias

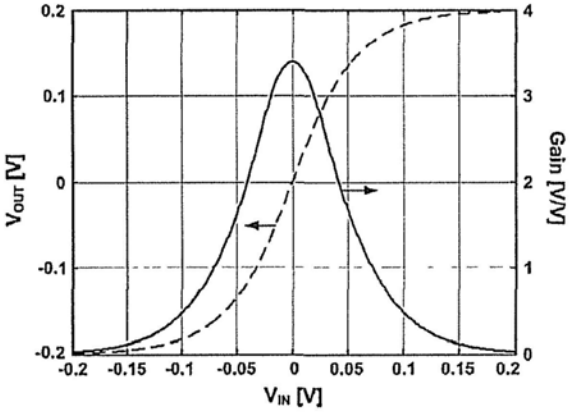


Figure 2.22 Voltage transfer curve and gain vs. input voltage

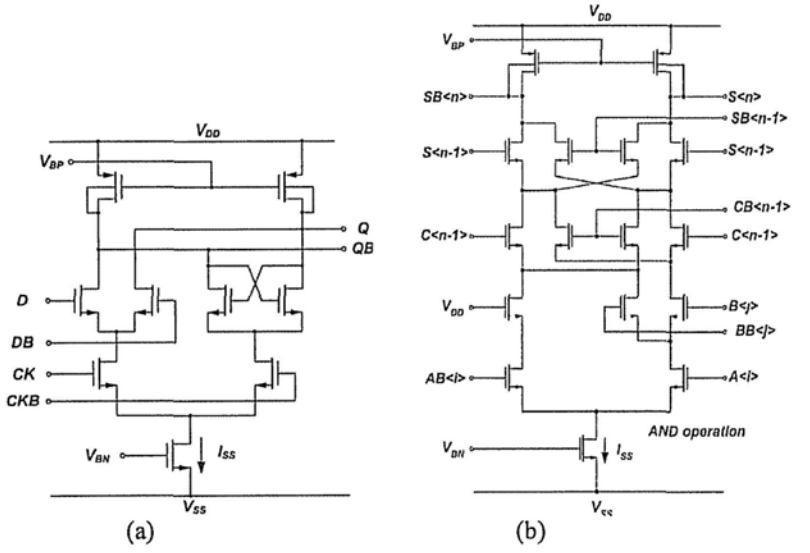


Figure 2.23 (a)STSCL D latch; (b)compound logic gate with the function of $\boldsymbol{S} < n >= S < n-1 > \oplus C < n-1 > \oplus (A < j > \cdot B < j >)$

applied in Figure 2.20. Figure 2.22 is the output transient curve and voltage gain versus the subthreshold input voltage. Figure 2.23(a) and (b) are the STSCL examples of D latch and a compound logic gate used in multiplier respectively [21].

STSCL has excellent structure in maintaining the stability of functionality in variations. But the topology may be complex and the trade-off between static surrent and propagation speed can be inefficient when the stacking of input network is large. The output level shifting is also another potential problem.

2.2.4 Small Scale Design

Based on the above logic styles and structures, some small designs for subthreshold operation were successfully implemented. Not only in the topology optimization in logic cells, the architecture also has influence on the performance of the design. In subthreshold region, logic delay is extended to the magnitude of 100 ns, parallelism or pipelined structure is required in most of the design. The leakage reduction is also a key method to guarantee the functionality.

Figure 2.24(a) is the conventional SRAM cells in read operation. When VDD is as low as 100 mV, the pull-down operation of SRAM cell #0 fails due to the leakages from 127 other cells which are connected to the same bit line $\overline{\text{RBL}}$. By using the small size multiplexer, the leakage problem can be solved in 100 mV read operation, which is shown in Figure 2.24(b) [17].

One of the basic arithmetic components in digital circuit is the multiplier. And a 8×8 multiplier was implemented in STSCL style in 180 nm technology [21], using the STSCL complex logic cell shown in Figure 2.23(b) as the array unit. The whole STSCL multiplier was compared with the CMOS multiplier in measurement. As what we can see from Figure 2.25, by tuning the tail bias current ISS, the STSCL

multiplier can operate in different frequency at the same voltage power supply. But it shows no improvement in power-frequency performance when compared to static CMOS style, and the cost of complexity is much higher than the static CMOS one.

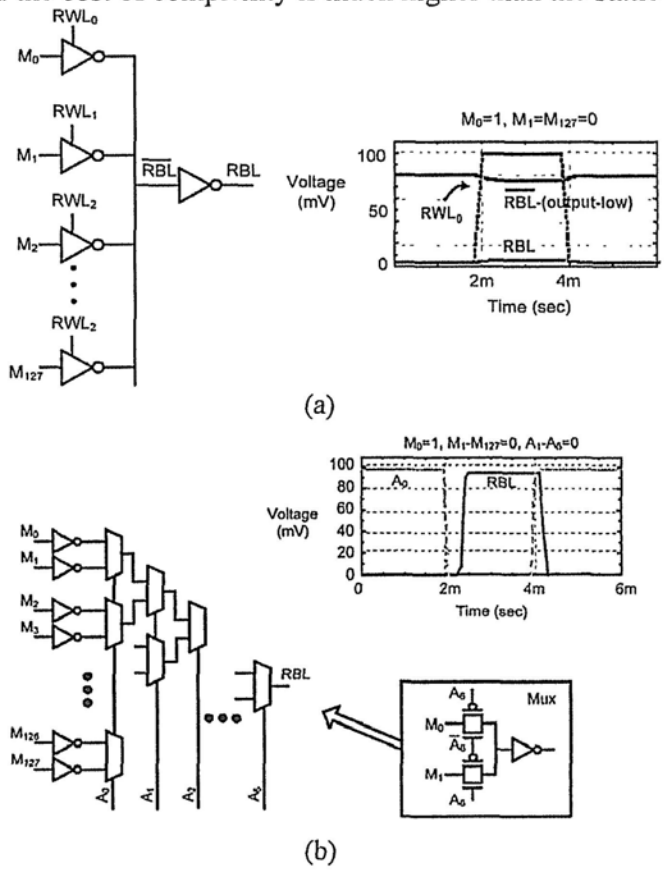


Figure 2.24 (a) Conventional SRAM array with leakage hazard; (b) leakage reduction SRAM topology for subthreshold read operation

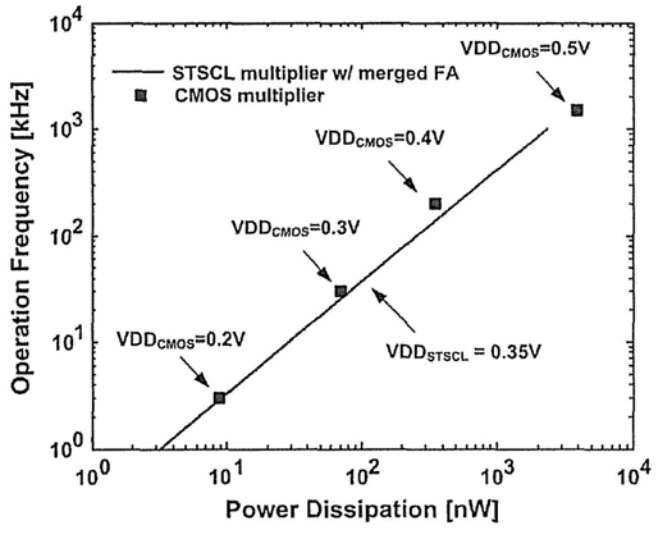


Figure 2.25 Comparison of CMOS and STSCL multiplier in power-frequency characteristic

There are also some optimized static CMOS designs for subthreshold operation. Like the FFT (fast Fourier transform) processor in [17], in which the optimized XOR gate and DFF in subchapter 2.2.1 are utilized. Other sizing and static topology optimized logic cells are also adopted in this design. Fabricated in a standard 0.18 um CMOS 6m dig tech. Occupies 2.6 X 2.1 mm, 627 000 transistors. The design is functional at 128, 256, 512, and 1024 FFT lengths, with 8-bit and 16-bit precision. The range of VDD is from 180 to 900 mV, featuring the maximum operating frequency from 164 Hz to 6 MHz. Figure 2.26 shows the energy consumption in every FFT operation versus VDD [15][16][17].

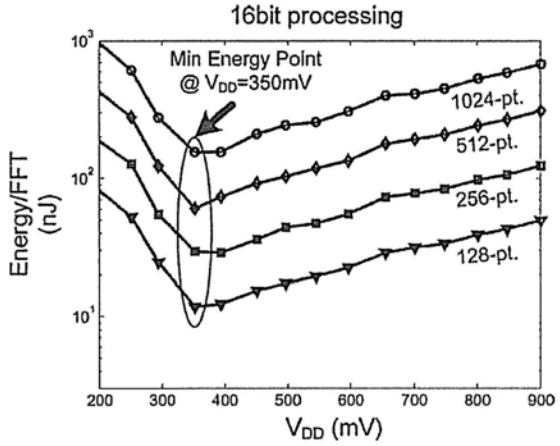


Figure 2.26 Energy per operation vs. power supply for the FFT processor

From Figure 2.26 we can recognize that the minimum energy consumption does not occur at the lowest VDD. This can be explained by the measured results from another design: a subthreshold CMOS 8-bit, 8-tap FIR filters. This design also adopts the sizing optimized CMOS logic cells. In worst-case condition, this design can operate in a lower voltage at 0.3 V, while the normal design fails to function below 0.4 V. As shown in Figure 2.27 [1][15], when VDD goes lower, the processing frequency of the filter reduces fast, and a smaller VDD gives smaller dynamic energy consumption (solid curves). Conversely, since the period for each processing cycle

becomes longer, the wasted leakage energy increases as it is proportional to the length of cycle. Once the decrement of the dynamic energy consumption can't compensate the increment of leakage energy, the total energy for each cycle or procession will increase. That is why the minimum energy consumption doesn't occur at the minimum operating voltage.

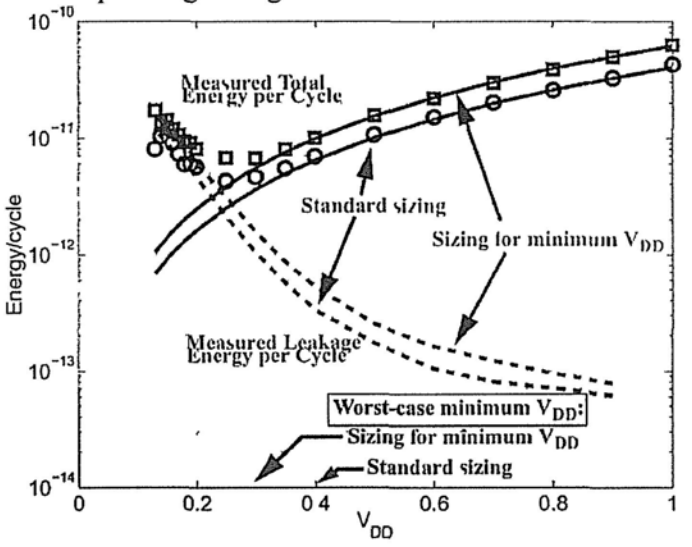


Figure 2.27 Energy per operation vs. power supply for the FIR filter, indicating the total energy consumption, and the dynamic and static (leakage) parts.

2.3 Characteristics of Subthreshold-voltage Devices and Logic in the Target Deep-Submicron Technology

2.3.1 Device Characteristics

In this thesis, all the designs are fabricated in UMC 130nm 8 metal-layer mix-mode process. To develop efficient logic cells and circuits in subthreshold operation, the electrical features of MOS transistors and the characterized data of general standard digital cells are essential. They help us to understand the how big the differences will be when the power supply scales down from nominal above 1 V to the level below 0.4 V. There are three device options in the target process: low leakage (LL) device with high threshold voltage Vth, standard performance (SP)

device with regular Vth and the high speed (HS) device with low Vth. The LL device has ultra low leakage but very weak current strength in subthreshold operation, conversely the HS device has a strong current strength but the OFF state leakage is relative high. The SP device has comparatively high driving current while the leakage current is much smaller than the HS device (see in Figure 2.28 (a)). Therefore in this design SP devices are selected for the balance between leakage and drive current. And generally the SP device has lower cost in large-number production.

The first thing to observe in the SP device is how weak the drain-source current can be in subthreshold region. And besides the exponential dependence on V_{GS}, how do other factors affect the amount of the current. Figure 2.28(b) and (c) are the drain-souce current versus gate-source voltage in different condition of power supply, they are for PMOS and NMOS transistor respectively. For each curve, the $|V_{GS}|$ varies from zero to the voltage level of the corresponding VDD.

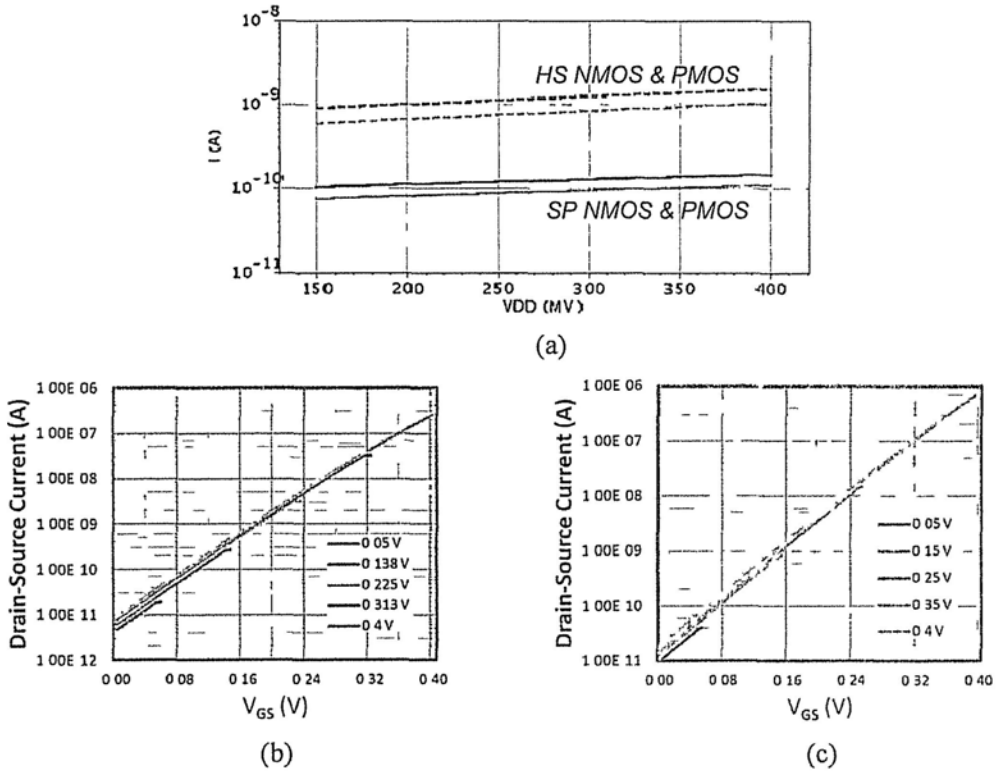


Figure 2.28 (a) leakage comparison of HS and SP device, Drain-source current vs VGS in different VDS: (b)320 nm-width SP PMOS: (c) 400 nm-width SP NMOS

However, merely this profile of I-V character could not reflect all the features of MOS transistors in subthreshold design. The drain-source voltage, the size of transistors and the body bias are all required in consideration.

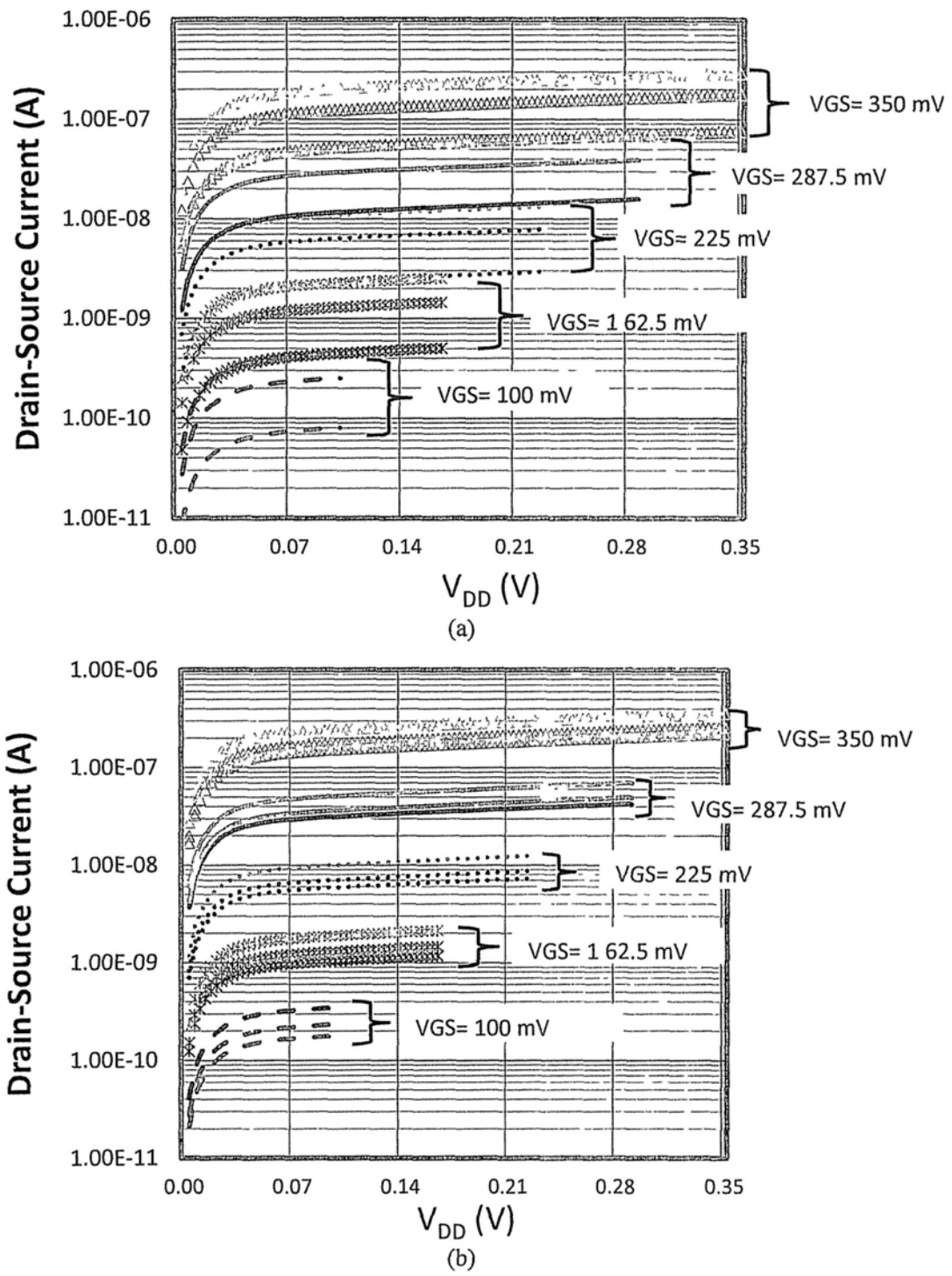
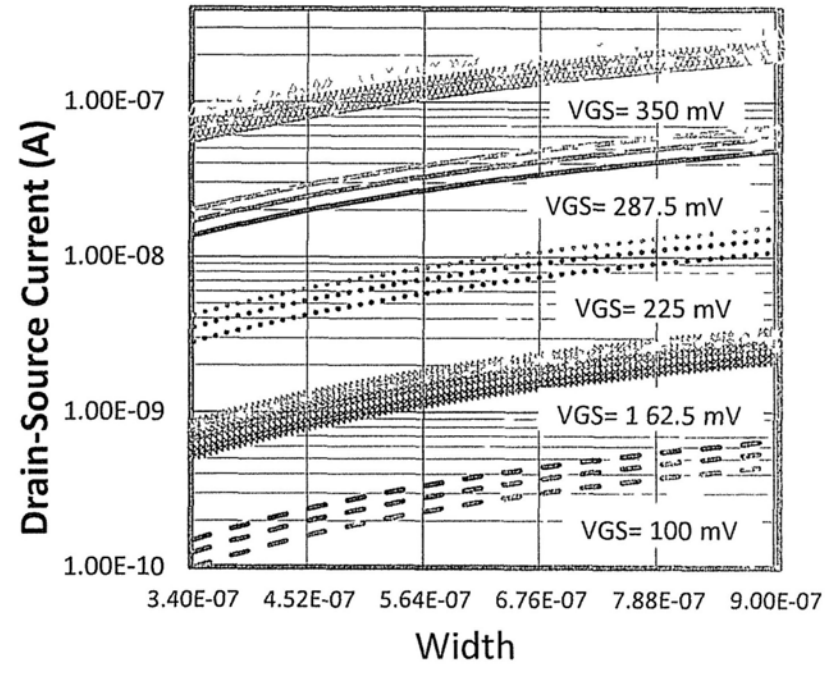


Figure 2.29 (a) PMOS Drain-source current vs. VDS in different VGS and width; (b)NMOS Drain-source current vs. VGS VDS in different VGS and width

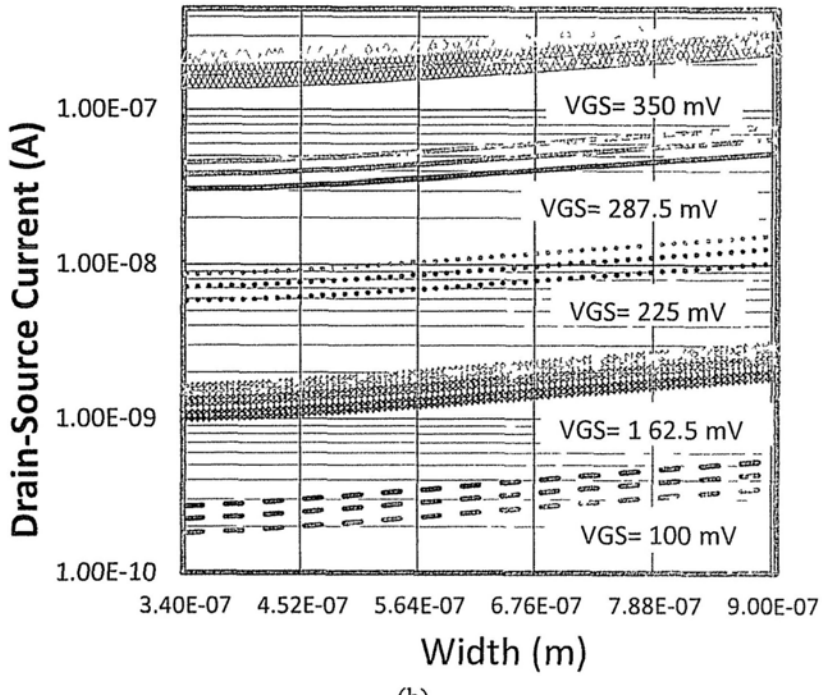
According to Equation (2.4), V_{DS} also has influence on the drain-source current of MOS transistors, when V_{DS} is closer the the thermal voltage V_T, the influence is more obvious. Figure 2.29(a) and (b) give another aspect of view in I-V curves, plus the options of V_{GS} and channel width (300 nm, 600nm and 900 nm for PMOS; 280 nm, 580 nm and 880 nm for NMOS). We can see that the influence from V_{DS} becomes slight as it goes further from the thermal voltage level. But compared to the condition in super-threshold region, the channel length modulation by V_{DS} is more obvious in effect, approximately 30% ~ 80% increment from 1/5 of the maximum value of V_{DS} to the full scale. In super-threshold region, such increment are generally less than 25%. The larger size of the width, results in more increment of drain-source current in percentage.

Refer to the sizing influence, because of the boundary effects, the drain-source current of PMOS and NMOS increase slowly with the width size when it is less than 500 nm. The slope of I-width curve becomes a bit steeper after 500 nm of the width size. Figure 2.30(a) and (b) gives the current-sizing curves of PMOS and NMOS transistors in different V_{GS} and V_{DS} (100 mV, 225 mV and 350 mV) condition.

In all of the current-related factors, the most interested one is the body-bias voltage. It is the most efficient way to adapt the threshold voltage of transistors as the design is already fabricated, giving the workable alternative method in ON-current control automation. The design in sub chapter 2.2.2 is an example. Given the exponential I-V dependence, a slight reduction of Vth can induce obvious current increment. However, in the target process, body bias for subthreshold operation may not be a good method. Figure 2.31 (a) and (b) are the current-bias curves in different V_{GS} (V_{DS} = V_{GS}) and width sizes (300 nm, 600 nm and 900 nm for PMOS; 340 nm, 620 nm and 900 nm for NMOS).



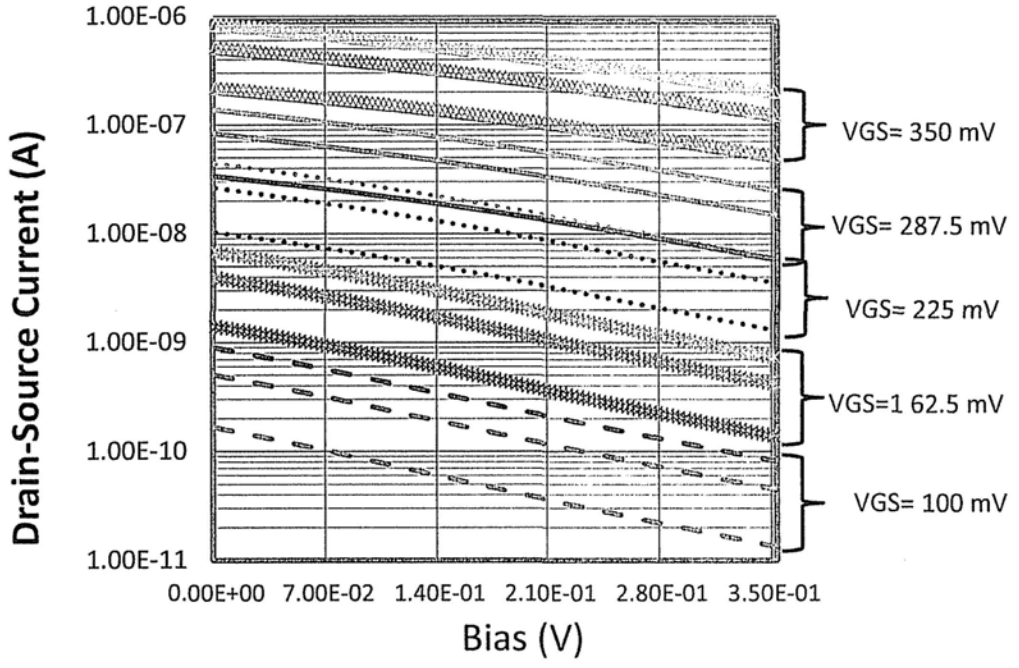
(a)



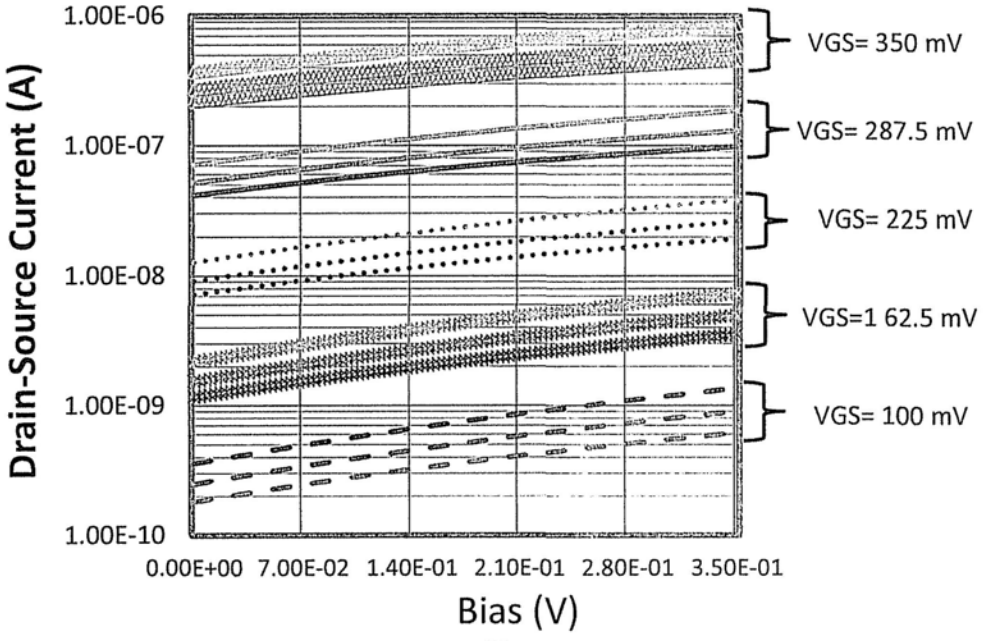
(b)

Figure 2.30 (a) PMOS Drain-source current vs. width in different VGS and VDS; (b)NMOS Drain-source current vs. width in different VGS and VDS

If body bias is just used for larger current, it will not have superiority than width extension. In Figure 2.31, we can see that even the width of zero-biased transistor is less than 3.5X width of the fully-biased one, its current is still much larger than the



(a)



(b)

Figure 2.31 (a) PMOS Drain-source current vs. body bias in different VGS and width; (b)NMOS Drain-source current vs. body bias in different VGS and width

biased one. If body bias is utilized in automatic ON current boost-up, the area penalty can be high. Due to the design rules, the NMOS device with body bias option requires the twin well technology. And the twin wells spacing must be longer than 1 μm if their internal NMOS transistors have different bias voltage. What's more, as

shown in Figure 2.32, the twin well edge must keep a 1.5 μm spacing to the edge of N well which enclosure twin wells with NMOS inside. Compared to a two-finger 1 μm NMOS in regular process, the 340 nm NMOS in twin well effectively occupies larger area, while their maximum current are almost the same. Besides the area penalty, the twin well technology requires a higher cost due to the more complicate process.

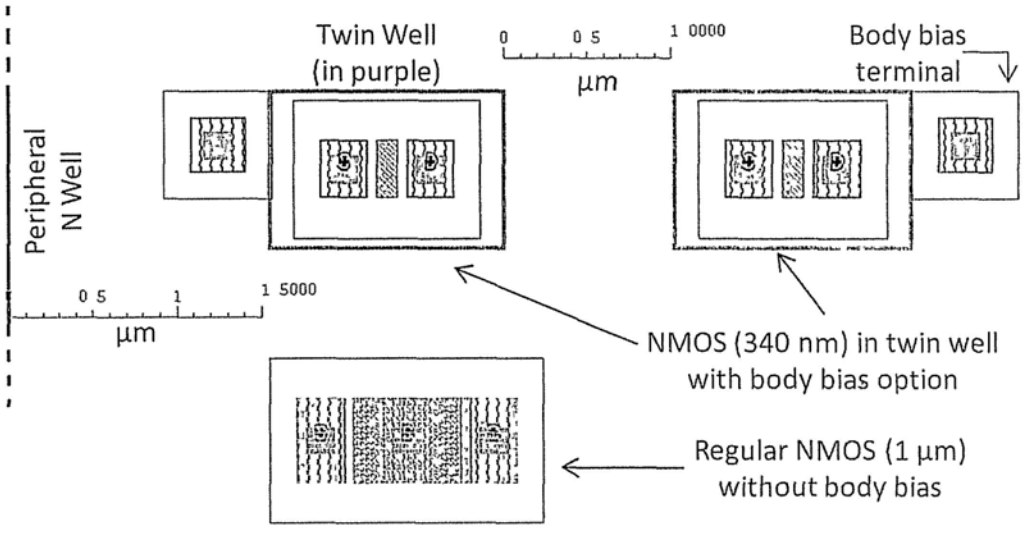


Figure 2.32 Area comparison of regular NMOS and body biased NMOS in twin well. They have nearly the same current driving strength

2.3.2 Standard Logic Cells Characteristics

Standard digital library gives convenience to digital designers in standard design flow. It removes the heavy work load of building up basic logic gate units from schematic to layout, providing abundant alternatives of logic cells with various functions. It also offers the accurate timing and power information to designers and EDA (electronic design automation) tools, helping them in architecture decision. However, this information is based on the nominal voltage operation. Industries seldom provide the subthreshold characterization data of the standard digital cells developed by them.

In large scale digital integrated circuit designs, design and verify all the logic cell units specified for subthreshold usage is a greatly heavy workload. Even the jobs of characterization require a quite long period. To implement a subthreshold digital design with efficiency of time and resources, it is much better to adopt most of the sources and steps in the general design flow. However, it is not completely as the same as the super-subthreshold design. According to the leakage current and transistor stacking principles in this chapter, high fan-in standard logic cells with over-three transistors stacking should be filtered out in the standard cells selection of the logic synthesis step, as well as the standard cells that may have OFF leakage hazard problem.

Besides the selection of proper standard logic cells for subthreshold digital design, understanding the timing and power specification is the key to a successful design that satisfies the frequency and functionality requirements in subthreshold region. An highlighted attention for subthreshold designers is the ultra great difference of logical delay from the one at nominal voltage operation. Extended from the magnitude of pico seconds to hundred nano seconds, the ultra weak current in subthreshold region causes the big change in timing.

To characterize the timing information of standard cell, related matrixes are required. Each element in a matrix represents the logical delay of the standard cell between one of the inputs and the output, while other inputs are kept in a certain pattern. The column and row position of the element represents the transition time of the related input and the capacitive load at the output. Like the standard cells in super threshold operation, input transition relates to how soon the ON current can reach at maximum and the amount of output load decides how long the ON branch of transistor can charge/discharge the node. These two factors affect the delay time of

logic cells, as well as the supply voltage levels (the number od levels and the value of them is determined by the target operating voltage range). Figure 2.33 is a brief demonstration of the characterization. With the data in the matrixes, EDA tools can calculate the very-close timing values by numerical method, for every gate or node in the circuit. Therefore when such matrixes are made, EDA tools can easily build up the digital circuits that satisfy specific timing constrains.

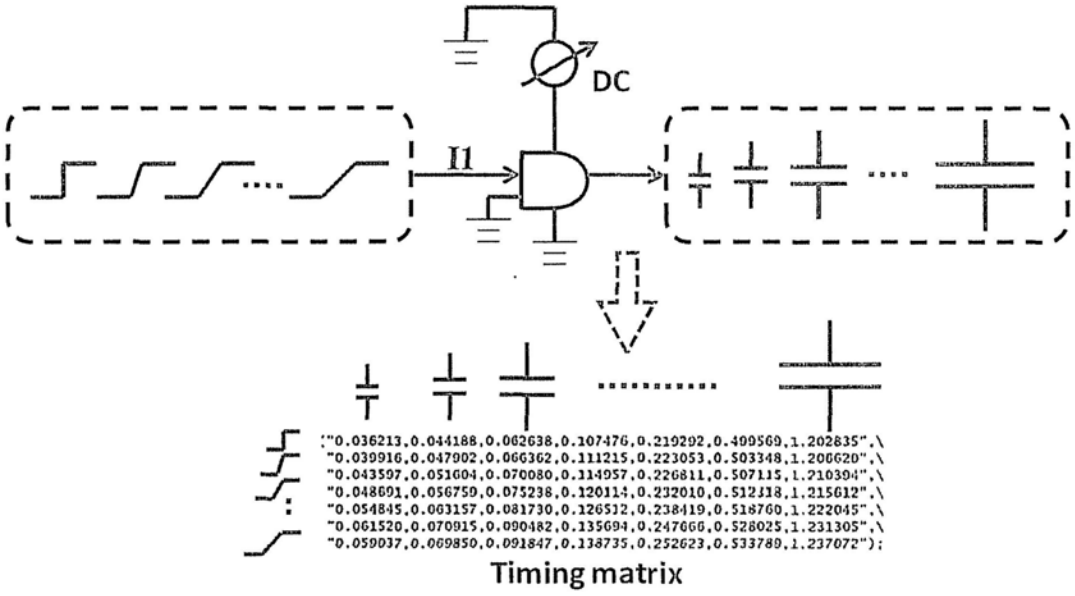


Figure 2.33 Example of the timing characterization of an AND gate. The output status is related to one of the inputs while the other is '0'. At a certain power supply voltage, process corner and in a specific input behavior, a related matrix will be made according to different input transition time and output capacitive load.

To obtain the timing and power information of standard cells in subthreshold information, several methods are available with different merits. The first one is to physically implement target standard cells on a chip in chain structures and with proper test circuits surrounded. Once the chip is fabricated, measurement can be done to characterized those cells with high and real accuracy. In this way it requires relatively longer period and heavier work load. A more efficient method is that the subthreshold standard cells can be characterized by simulation. Given the detail

layouts of the cells and the process modeling files are accessible, we can use EDA tools to extract the parametric circuit netlists and simulate their responses with certain input patterns. The spent time and consumed resources of this one are much less than the first method, even though a lot of computer jobs are necessary. However, it is also with less accuracy than the first one. If the entire design or some blocks of it are not with critical timing requirement, the third method is nearly the easiest way to run characterization. Or it should be called estimation, which is more appropriate. The whole estimation is based on a famous equation (Equation 2.13), which describes the scaling of t_d, the gate circuit's logical delay, as VDD is changed. K is a delay fitting parameter, C_L is the effective output load capacitance, and α is a process dependent parameter. If the .lib files of standard cells library are available, which contain the timing information of cells operating at nominal voltage range, we can use Equation (2.13) to calculate the approximated mapping values when VDD scales down to subthreshold region. In Equation (2.13), t_d is the logic delay at a certain supply voltage V_{DD}. Given the power supply V_{DD} changed from nominal 1.8V to 0.4V, then by calculation the logic delay $t_{d,\,0.4V}$ at 0.4V V_{DD} is 9.87 times the $t_{d,\,1.8V}$ at 1.8V V_{DD}. (Here α is 1.3 and V_T is 0.32V in 180nm CMOS technologies). That means if a design satisfies the clock frequency of $9.87\times(t_{d,\,0.4V})^{-1}$ at 1.8V supply, approximately it can satisfy the clock frequency of $(t_{d,\,0.4V})^{-1}$ at 0.4V supply. This mapping relationship is used to ensure the operating range.

In this way the characterization job is greatly reduced but the accuracy is lack of guarantee. It must give enough timing margin in the design to deal with uncertainty. The above three methods can be shown in flow chart format in Figure 2.35. Figure 2.34 is a brief comparison of these three methods.

$$t_d = K \frac{C_L V_{DD}}{(V_{DD} - V_{th})^{\alpha}} \tag{2.13}$$

Figure 2.36 gives several examples of frequently used standard logic cells from a library based on the target technology. They are the timing characterization with different input transition and output capacitive load. The selection of the transition times and capacitive values are featured ones of the library in subthreshold operation. The delay times are over 100 ns as the operating voltage is 0.25 V, reflecting the subthreshold timing magnitude.

More detail of timing characterization of other standard or custom cells, and explanations will be given in following chapters. The examples here are for the indication of subthreshold timing profile.

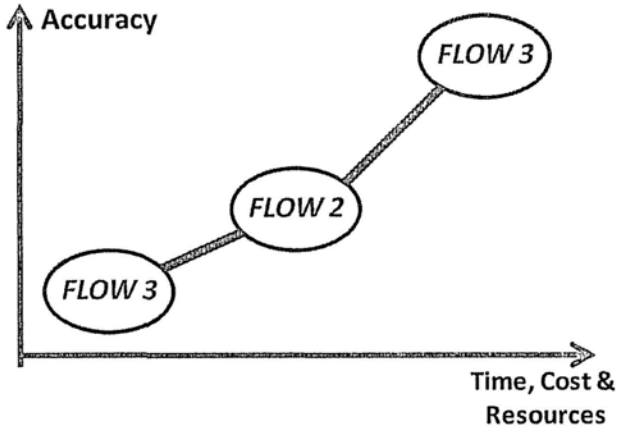


Figure 2.34 Comparison of the three methods of subthreshold characterization

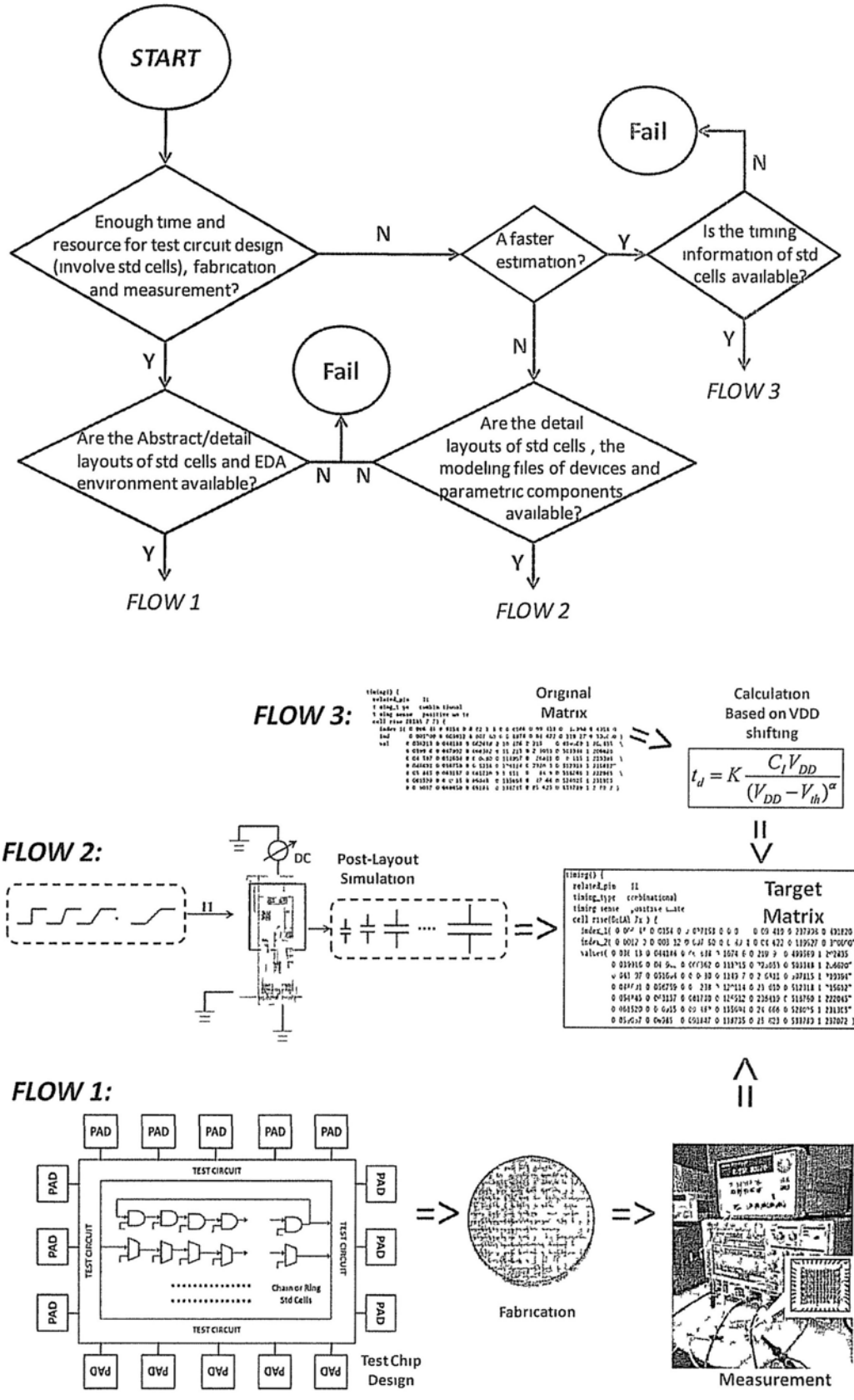
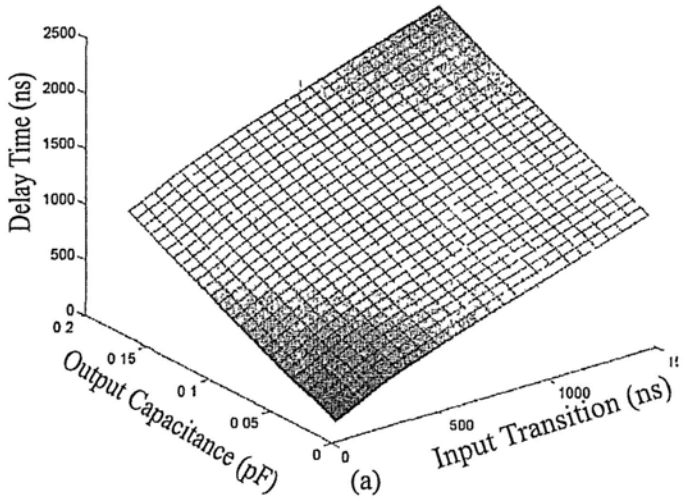
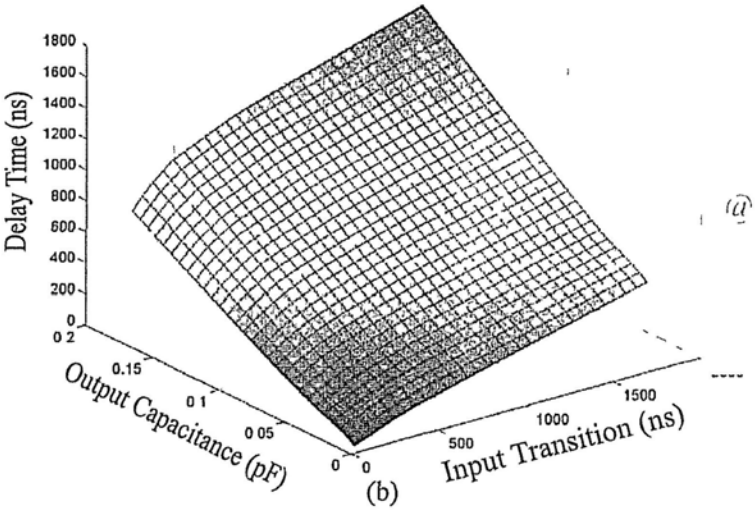


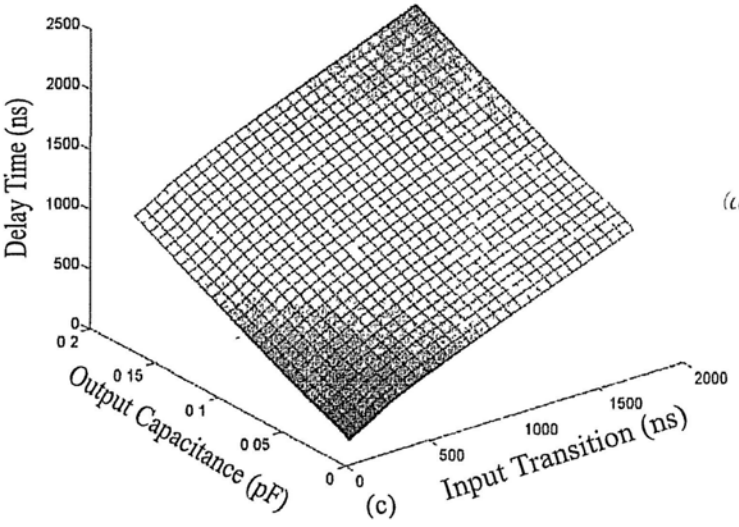
Figure 2.35 Methods and flows for subthreshold logic cells' characterization



(a) 0.25 V, 25 °C, TT proc. corner



(a) 0.25 V, 25 °C, TT proc. corner



(a) 0.25 V, 25 °C, TT proc. corner

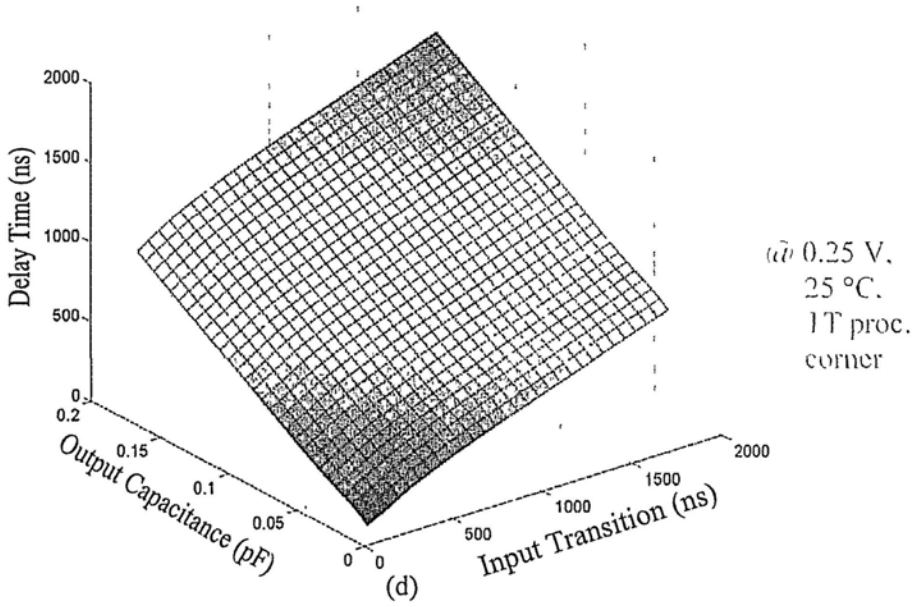


Figure 2.36 (a) 2-input XOR, matrix of the delay from input I2 to output, when I1='1', I2 falls from VDD to ground. (b) 2-input NAND, matrix of the delay from input I2 to output, when I1='1', I2 rises from ground to VDD. (c) D Filp Flop, matrix of the delay from input CK to output Q, when input D='1', CK rises from ground to VDD as triggered edge. (d) 1-bit Full Adder, matrix of the delay from input CI to output CO, when A='1', B='0', CI rises from ground to VDD.

CHAPTER 3 BASEBAND PROCESSOR IN PASSIVE RFID TAG

3.1 Introduction

Radio frequency identification (RFID) is an automatic wireless data collection technology with a long history First functional passive UHF RFID systems with a range of several meters appeared in early 1970's [33] Since then, RFID has experienced a tremendous growth RFID bands vary in different countries and include frequencies from 125 kHz and 13.56 MHz (ISO/IEC 14443, 18000-3 and 15693 etc), to the UHF region 433 MHz, 860 MHz and 960 MHz, until 2 4 GHz (EPCglobal standard, ISO/IEC 18000-6 and ISO 18185 etc)

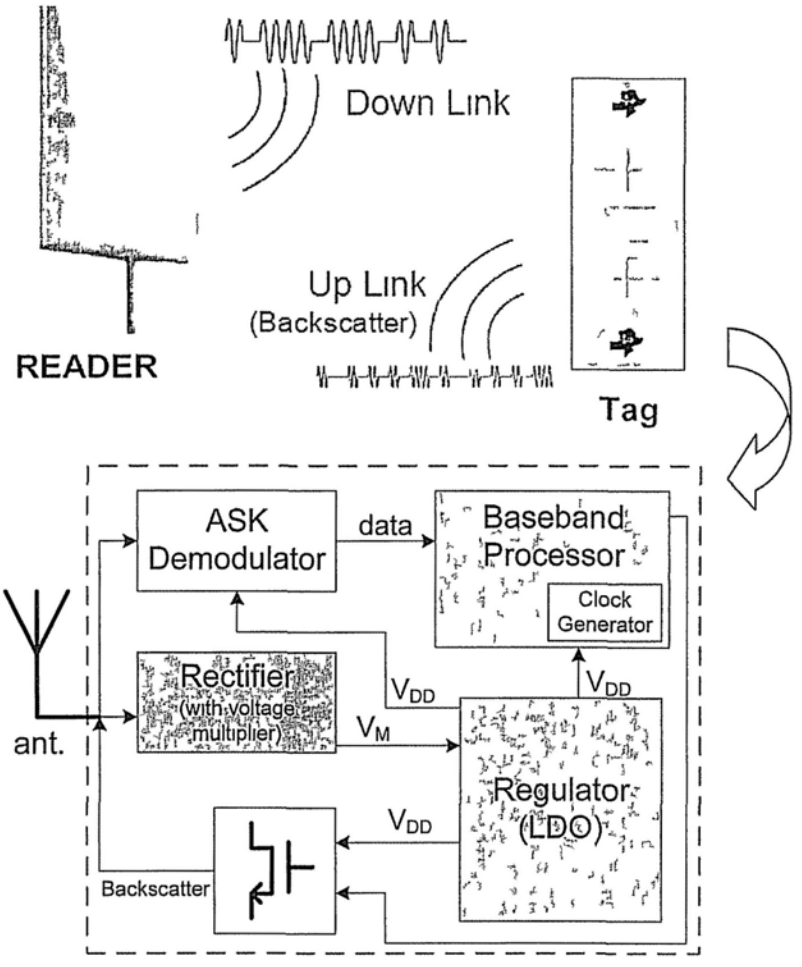


Figure 3 1 Reader and passive RFID tag with block structure

As shown in Figure 3.1, passive RFID tag utilizes antenna to collect energy from the reader-transmitted electromagnetic wave. The voltage rectifier in the tag roughly process the sine voltage wave from antenna, transferring the AC voltage to multiplied DC output, the following voltage regulator (low dropout regulator, LDO) helps to provide a stable DC power supply to the functional blocks: input-signal demodulator and, besides these physical layer components, the baseband processor (BBP). It is also the core of RFID tag. It takes charge of received signal decoding, processing and the related backscatter response. With the control of baseband processor, a communication handshake with passive RFID system operates in the following steps:

(1) RFID reader transmits a modulated RF signal to the RFID tag consisting of an antenna and an integrated circuit chip.

(2) The chip receives power from the antenna and responds by varying its input impedance to antenna and thus modulating the backscattered signal. Modulation type often used in RFID is amplitude shift keying (ASK) where the chip impedance switches between two states: one is matched to the antenna (chip collects power in that state) and another one is strongly mismatched.

(3) Shunting the antenna causes a momentary fluctuation (dampening) of the carrier wave, which is seen as a slight change in amplitude of the carrier. The reader peak-detects the amplitude-modulated data and processes the resulting bitstream according to the encoding and data modulation methods used.

Figure 3.2 demonstrate that how reader and tag communicate in a free space. The most important RFID system performance characteristic is tag range – the maximum distance at which RFID reader can either read or write information to the tag. Tag range is defined with respect to a certain read/write rate (percentage of successful reads/writes) which varies with a distance and depends on RFID reader

characteristics and propagation environment [35][36]. A typical inter-dependence of distance, frequency, and read rate is shown in Figure 3.2 for an RFID system consisting of a reader and a tag tuned to 900 MHz in a specific environment [35].

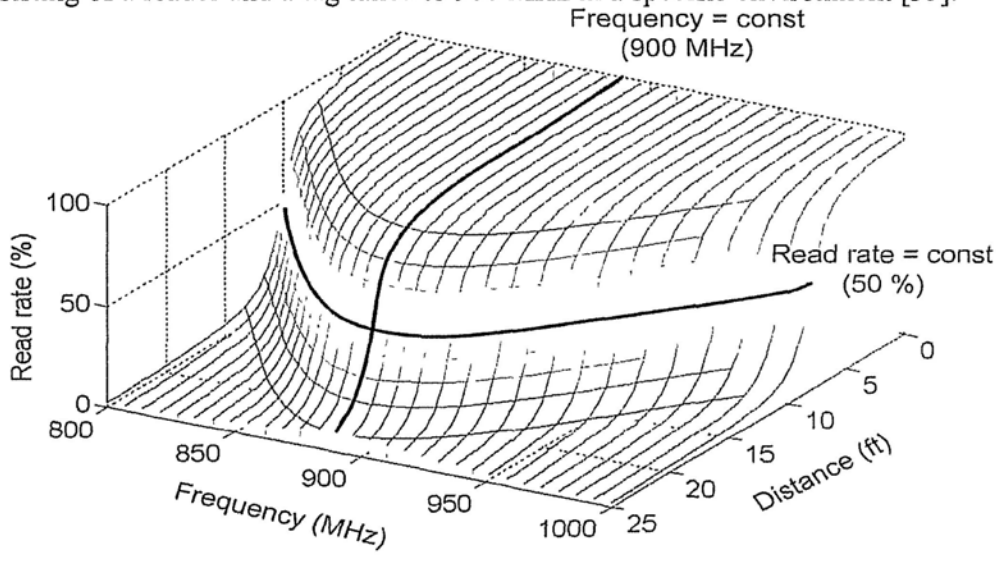


Figure 3.2 Read rate vs. distance and frequency in a typical RFID system.

As a passive system powered by harvesting electromagnetic wave, the energy or power obtained from antenna is strongly influenced by the distance from the signal transmitter: the reader. The power of the transmitted signal is determined by a parameter called EIRP (equivalent isotropic radiated power). Further form the reader, weaker the signal power available. Figure 3.3 shows the signal power transmitted by the reader and backscattered from the chip of an ordinary tag, versus the distance from the origins, considering reader has an EIRP of 4 W (36 dBm) [35].

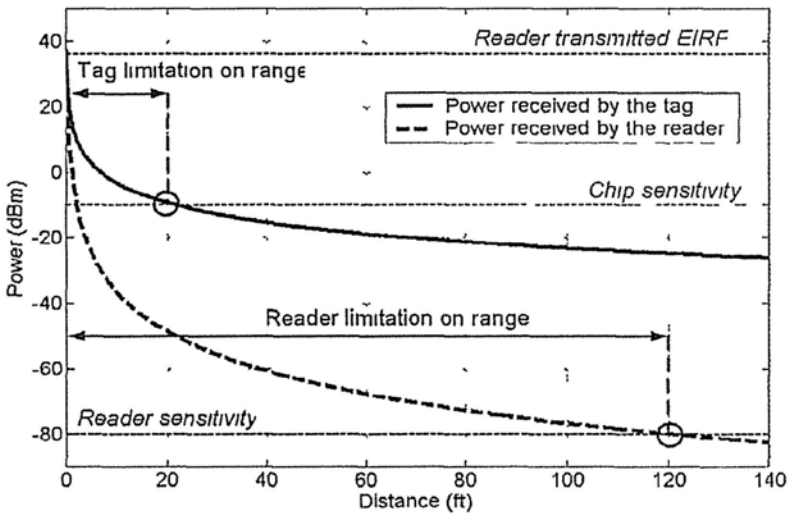


Figure 3.3 Received power vs. distance for tag and reader in RFID system

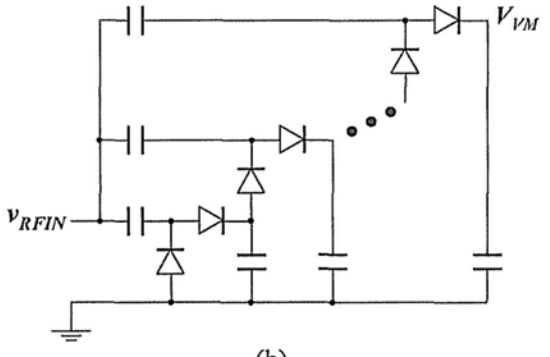
From Figure 3.3 indicates a general feature of passive RFID tag, the workable range of such RFID system in free space is defined by the limitation imposed by the tag (20 ft), not by the reader (120 ft). So chip sensitivity is the most important tag limitation. It is the minimum antenna received signal power necessary to turn on RFID chip. The lower it is the longer is the distance at which the tag can be detected [36][37][38]. As the baseband processor is taking a great part of power consumption, its power supply is somehow a key factor to the chip sensitivity. Scaling down the power supply to subthreshold region is absolutely with weak current operation. But since the baseband domain in RFID tag is within the range from low frequency to sub high frequency (less than 10 MHz), a proper digital design tailored to subthreshold region can probably overcome the shortcut brought by weak current in ON status. Once the drawback problem is solved, the merits of subthreshold operation are relatively more beneficial:

(1) Subthreshold voltage directly leads to ultra low power consumption based on the Equation (2.1) for dynamic power.

(2) As the supply voltage is greatly reduced, somehow the LDO module can be less delicate to maintain a high output level. The VDD for LDO is lower as well as its power consumption. Furthermore, when workable supply voltage of BBP is in the range of 0.15 ~ 0.3 V, which is close to the output of the rectifier, the general LDO can be substituted by a simple low pass filter. It is competent with much less complexity and power.

(3) The subthreshold operate-able BBP can also reduce the complexity and power of the rectifier. Since the output voltage does not need to boost up to a high level, as shown in Figure 3.4, it can reduce the number of series stage for voltage multiplication (if the tag is not required to contain a writable non-volatile memory) [40][41]. In this way, less stages, less power dissipation, and the conduction loss will decrease while the power efficiency will increase.

(a)



(b)

Figure 3.4 Rectifier (a) a diode-connected NMOS functioned as a diode (b) the structure of the rectifier with series stages.

The rectifier and regulator are analog blocks and always occupy a large silicon area. If they can be removed or reduced in size, it will be a great help on the cost decrease of the whole tag. Additionally, a subthreshold operating BBP does not only reduce the power itself, but also the power of the two analog blocks. Consequently,

the tag-reading constrain in Figure 3.3 can be decreased that the successful range of reader-tag signaling can be extended longer. Figure 3.5 is the measured output voltage versus the input power from the antenna of an 8-stage rectifier [40][41]. Combined with Figure 3.3, if the baseband processor can operate at 0.15 ~ 0.3 V in low power (with compatible demodulator, regulator and rectifier in tag), then the whole tag can operate in the range of 40 ~ 60 feet (12.2 ~ 18.3 m), twice or three times of the ordinary designs. In this work, the target subthreshold design example is the BBP of the passive RFID tag in UHF band.

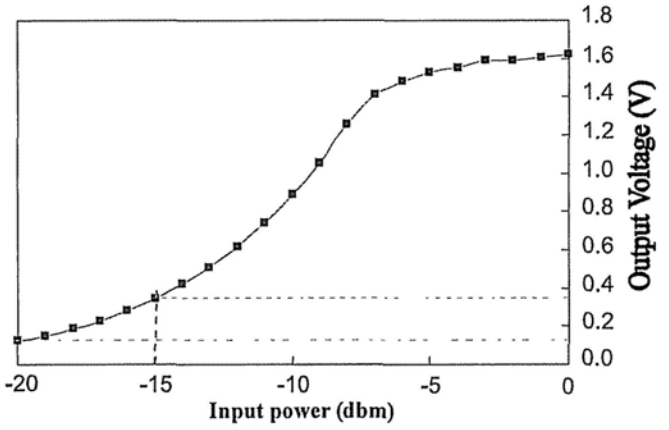


Figure 3.5 Measured output voltage vs. input power of 8-stage rectifier

3.2 Air Interface Communication and Protocol

3.2.1 General Frequency Characteristics in RFID Tags

Air interface communication is a wide conception including all the wireless communication behavior in a board sense. In RFID aspect, it is embodied in a series of protocols that specify the signaling rules, carrier frequency band, electrical parameters and data rates.

As shown in Figure 3.6 [32], the radio frequency has a wide range spectrum, from 100 k Hz to 10 GHz. Due to the frequency characteristics of electromagnetic wave in air interface, passive RFID tags with different featured frequencies have different usage according to their data rates and operating length.

Passive RFID devices operating below 135 kHz are merely used in item management applications. They have lower data rate and fewer channels, supporting only one reader environment. Our target baseband processor is aimed to apply in the RFID tags that based on the EPC Radio-Frequency Identity Class-1 Generation-2 Protocol for Communications at 860 MHz ~ 960 MHz. This frequency band is defined as UHF in the spectrum with greatly more channels. And the data rates are much higher as the carrier is now above 850 MHz. Refer to the operating range in air, UHF passive systems use the Electric field and transfer power by capacitive coupling, achieving greater reading ranges than LF and HF Tags which use the magnetic field.

With the features of higher, faster and longer, UHF passive tags are utilized in massively distributed sensor networks, airport baggage logistics and supply chain item-level tagging with enhanced security feature etc. And UHF devices support multi-reader, multi-tag environment. In a word, the capability of UHF tags is much more superior. Table 3.1 is the global frequency regulation of UHF tags [32]

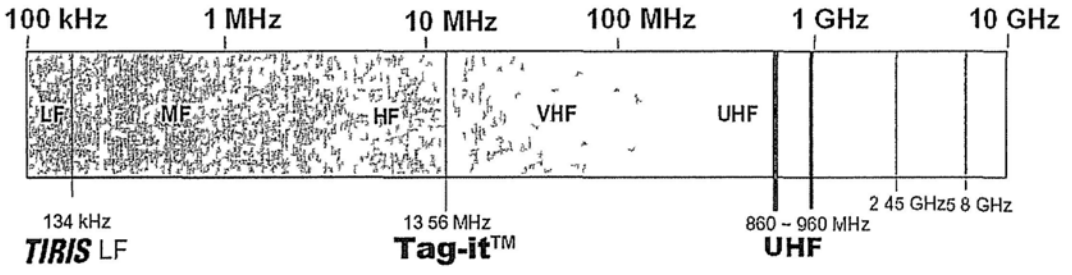


Figure 3 6 Radio frequency spectrum

Table 3.1 Global regulation of UHF tags

	North America	Europe	Japan	Korea	Australia	Argentina, Brasil, Peru	New Zealand
Band (MHz)	902~928	866~868	952~954	908.5~914	918~928	902~928	864~929
Power	4W EIRP	2W ERP	4W EIRP	2W ERP	4W EIRP	4W EIRP	0.5~4W EIRP
Number of Channels	50	10	TBD	20	16	50	Varies
Spurious Limits	-50 dBc	-63 dBc	-61 dBc	-36 dBc	-50 dBc	?	?

3.2.2 Generality of Communication and Specialty in EPC C1G2 Protocol

Besides the physical layer of power and signal transfer issues, the general pattern of the talks between readers and tags are the most important to ensure the efficiency of communication.

Two principles mainly regulate the communication. The first one, also being as part of the European regulations, is LBT (listen before talk). If a reader detects a signal in its environment, on the channel it intends transmitting, it must switch to another free channel. After 4 seconds it must turn its transmitter off for 0.1 seconds to allow other readers access to that channel.

The second one is RTF (reader talks first principle). For operation, a reader with antenna is required to send a command to the transponder (tag) and to receive its response (see Figure 3.1). The tag does not transmit data until the reader sends a valid request, that is a command with correct parameters to the tag. And then starts the communication round.

Reader-tag communication should obey the half duplex operation. That means both parties should wait for the response from the other after they have sent a message. This flow is general but not strictly constrained. Due to the quick-change

air interface and power supply, tag may miss the command from the reader or failed to response to a correct/wrong command. Therefore most tags' baseband processors have state machine scheme to cooperate with the state machine of the readers, keeping the communication away from dead end.

The state machine is a controller where groups of commands become available once the chip has transitioned to the correct state. Most communication protocols specify the tag to go through the similar state iteration like "Ready-Stand by-Reply-Ready". Depends on the complexity of the function and signal steps, the scale of the state machine can be larger or smaller. For the target protocol EPC C1G2 for UHF RFID, the state transition and flows are described in detail as following.

As soon as the chip enters the RF field it changes to the *Ready* state and will accept *Select* commands. *Select* commands are sent to all tags to inform each one if it is to take part in the *Inventory* process that is to follow. Multiple *Select* commands can be used to precisely define which tags are to respond. All exchanges among readers and tags start with one or more *Select* commands. Tags don't respond (backscatter) to the *Select* command.

Now the "*Inventory*" group of commands can be used to initiate the singulation process, where each individual tag is identified and processed. Each *Inventory* round starts with a *Query* command being broadcast - this command passes a Q-value (0 to 15) from which each tag generates a random slot counter number in the range (0, 2^Q - 1). Most readers dynamically adjust the Q-value depending on the number of tag in the field, thereby increasing the potential reading rate.

If a tag generates a slot counter value of zero it is allowed to reply by sending a 16-bit random number and at the same time transitioning to the *Reply* state. The other tag change state to *Arbitrate* and wait for further commands. If the tag's response is

successfully received, the Reader replies by sending an *ACK* command, together with the same 16-bit random number. This response now allows the tag to send back its *EPC* data (stored in the memory) and change state to *Acknowledged.*

It is at this point that the Reader is able to transition the tag to the *Open* (or *Secured*) state allowing operations such as *Read*, *Write*, *Lock* and *KILL*, but normally this exchange would terminate when the reader sends a *QueryAdjust* command and the tag switches state back to the *Ready* state and changes its inventoried flag to show it has been singulated. The *QueryAdjust* command also affects the other tag's causing them to decrement their slot counters and any tag whose counter is now zero is allowed to reply – so in this way with successive *QueryAdjust* or *QueryRep* commands all tag's will be found and singulated.

If two tag's reply at the same time, unless the Reader is able to identify each one and send an *ACK* and the correct 16-bit random number, each one will timeout, re-generate a slot counter value and return to *Arbitrate*.

If further actions need to be performed on a tag, once the tag has replied its EPC number and is in the *Acknowledged* state, the Reader sends a *Req_RN* (Request Random Number) command. The tag replies with a new 16-bit random number that is called the *Handle* and changes its state to *Open*. The Handle then becomes the token for further commands such as Reading, Writing or Killing. If the *Lock* command is required and the access password is non-zero, the *Access* command (complete with Access Password) must be sent, to cause the tag to change to the *Secured* state. Once in *Secured* state, all commands are available.

Figure 3.7 give the basic state transition in *Inventory* process, which is the key process period of the whole communication round [42].

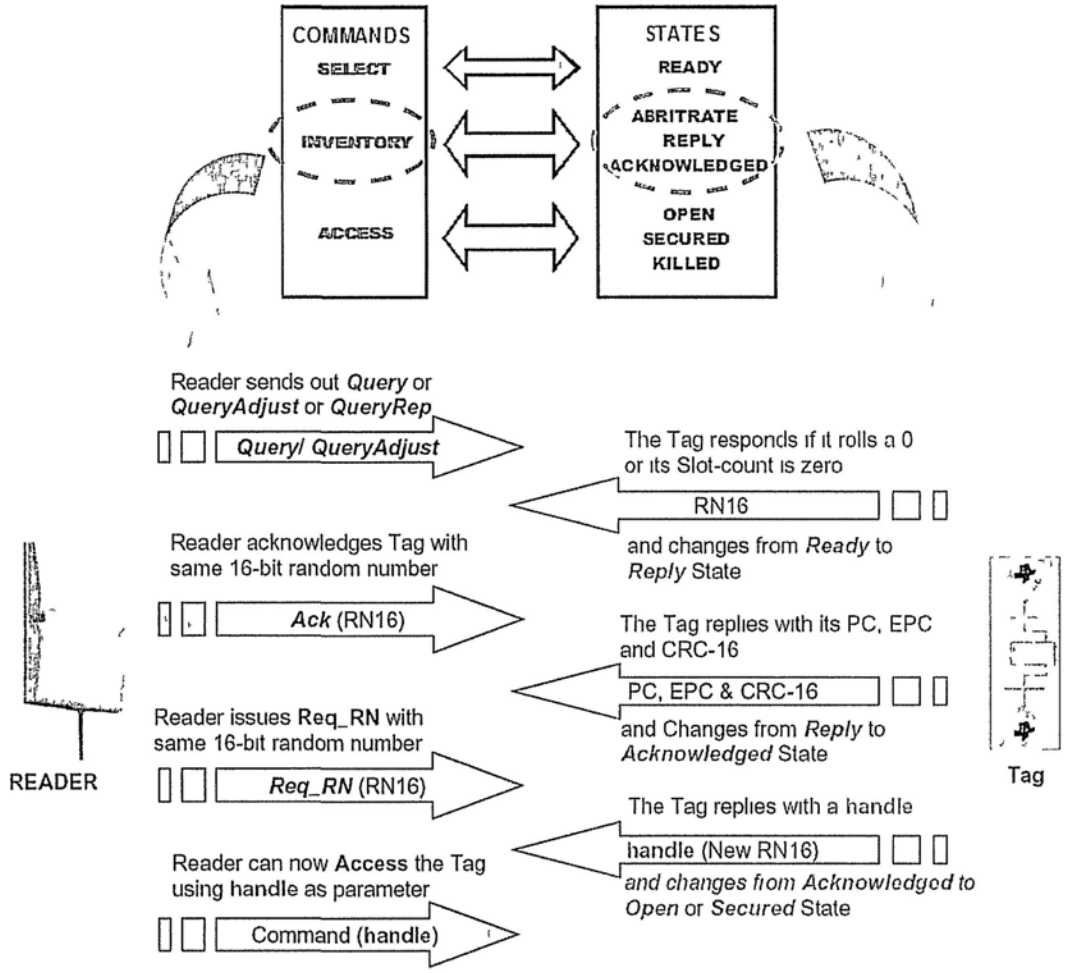


Figure 3.7 Commands and state transition between reader and tag

More details of state transition and commands description of EPC C1G2 protocol can be seen in Appendix A and B. The parameters PC and EPC are information of the article stuck with the tag. They are stored in the BBP-external memory banks.

3.3 Systematical BBP Design based on EPC C1G2 Protocol

3.3.1 Data Link Portion

The baseband processor that based on EPC C1G2 protocol is the target subthreshold digital design of this thesis. The target BBP design covers most of the commands in the categories of Select, Inventory and Access. It also covers most of the states specified in the protocol. Besides the core processing component of

baseband processor, the data link portion and anti-collision blocks are also involved in the BBP design.

EPC C1G2 UHF RFID communication scheme uses PIE (Pulse interval encoding) format in reader-to-tag ASK (Amplitude Shift Keying) modulation. PIE format defines the symbols of data-1 and data-0 according to their length of interval. And in reader-to-tag RFID communication, it also has several specific timing reference symbols, for the tag to calculate the intermediate threshold interval between data-1 and data-0; and to calculate the interval parameter in backscattering, using another decoding format. There are two parts in a PIE symbol: the high level part with variable duration and the low level part with relatively fixed duration called *PW*. When the reader transmits signal to tags, these two parts stands for two different modulation depths.

As shown in Figure 3.8, *Tari* is the reference time interval for Interrogator-to-Tag signaling, and is the duration of a data-0. In the protocol *Tari* can be any value between 6.25 μs and 25 μs. Readers shall use a fixed modulation depth, rise time, fall time, *PW*, and *Tari* for the duration of an inventory round.

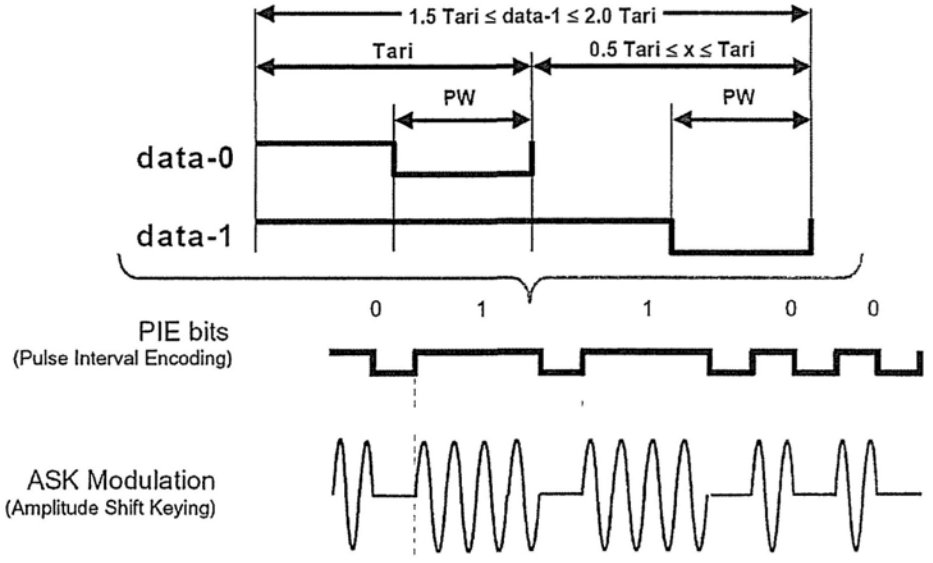


Figure 3.8 Interval of data symbols in PIE format

Table 3.2 lists parameters specification of the RF envelop for reader to transmit modulated signal. When the reader transmits enveloped RF signal, the whole packet will be sent serially symbol by symbol, header goes first [43]. The header includes PIE symbols of timing reference, and it is dependent, as shown in Figure 3.9.

Table 3.2 RF envelope parameters

Tari	Parameter	Symbol	Minimum	Typical	Maximum	Units
6.25 µs to 25 µs	Modulation Depth	(A–B)/A	80	90	100	%
	RF Envelope Ripple	$M_h = M_l$	0		0.05(A–B)	V/m
	RF Envelope Rise Time	$t_{r,10-90\%}$	0		0.33Tari	µs
	RF Envelope Fall Time	$t_{f,10-90\%}$	0		0.33Tari	µs
	RF Pulsewidth	PW	MAX(0.265Tari, 2)		0.525Tari	µs

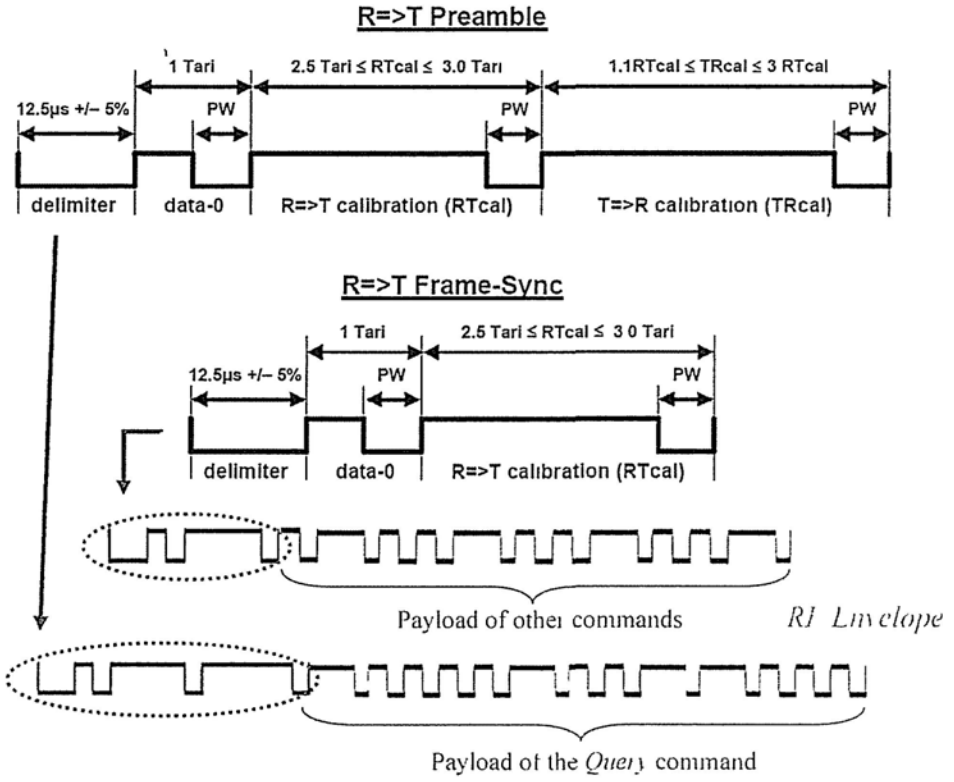


Figure 3.9 Header of reader-to-tag signal packets

A reader shall set *RTcal* equal to the length of a data-0 symbol plus the length of a data-1 symbol (*RTcal* = *0length* + *1length*). A Tag shall measure the length of *RTcal* and compute *pivot* = RTcal / 2. The Tag shall interpret subsequent reader symbols

shorter than *pivot* to be data-0s, and subsequent reader symbols longer than *pivot* to be data-1s. The Tag shall interpret symbols longer than 4 *RTcal* to be bad data. Prior to changing *RTcal*, an reader shall transmit CW for a minimum of 8 *RTcal*.

The timing reference *TRCal* in the preamble is not for PIE decoding but backscatter encoding. The tag uses FM0 or Miller encoding to backscatter signals to the reader. And the backscatter link frequency (*LF*, its FM0 data rate or the frequency of its Miller subcarrier) using the *TRcal* and divide ratio (*DR*) in the preamble and payload, respectively, of a *Query* command that initiates an inventory round. Equation (3.1) specifies the relationship between the *LF*, *TRcal*, and *DR*. A Tag shall measure the length of *TRcal*, compute *LF*, and adjust its T=>R link rate to be *equal* to LF (40~640 kHz, Chapter 4 will give *LF* details and tolerances). Detail content of the subthreshold backscattering encoder design is also given in Chapter 4.

$$LF = \frac{DR}{TRCal} \tag{3.1}$$

In data link between reader and tag, there is timing requirement that regulates the half duplex communication. Details are given in Appendix C.

To facilitate the anti-collision and to avoid signaling mistakes, random number and redundancy check are necessary in the BBP. EPC C1G2 protocol uses CRC-16 and CRC-5 for cyclic-redundancy check, the number stands for the bit size. CRC-5 is applied in the redundancy check of incoming *Query* command. CRC-16 is a frequently-used check that a reader uses when protecting certain R=>T commands, and a tag uses when protecting certain backscattered T=>R sequences (see the CRC position in packets from Appendix A and B). To generate a CRC-16 or CRC-5, an reader or Tag shall first generate the precursor, then take the ones-complement of the generated precursor to form the CRC. Table 3.3 is the CRC definition.

Table 3.3 CRC definition

CRC-5 Definition				
CRC Type	Length	Polynomial	Preset	Residue
—	5 bits	$x^5 + x^3 + 1$	01001_2	00000_2

CRC-16 precursor				
CRC Type	Length	Polynomial	Preset	Residue
ISO/IEC 13239	16 bits	$x^{16} + x^{12} + x^5 + 1$	$FFFF_h$	$1D0F_h$

3.3.2 Functional Partition

According to the protocol specification, the baseband processor design can be structurally divided in several ways. Depends on the period of processing, the BBP can be divided into blocks of receiver, handler, and transmitter. Depends on the state transition diagram, the BBP can be divided into external data link interface, inventory handler, access handler and memory interface. There are also other methods to partition, the selection must be based on the special requirement.

The focus of this thesis is on the low voltage and low power, especially at the ultra low voltage to subthreshold level. Ultra tense timing condition is the primary concern when drawing the block diagram. According to the protocol, the whole baseband processor works in with two clock signals. Therefore, the initial partition is based on the clock domain.

As shown in Figure 3.10, the whole design is roughly divided into two main parts with connections. The original clock source is from the *Ring Oscillator* (RO), which can be considered as the small analog part of the BBP. The original clock signal helps to count the length of data symbols in the RF envelope from the demodulator. And its period is also the minimum unit to be counted to generate the clock of link frequency. The *Command Handler* here is take charge of state transition and main

control, and is driven by the clock for LF, making it synchronize with the backscattering steps.

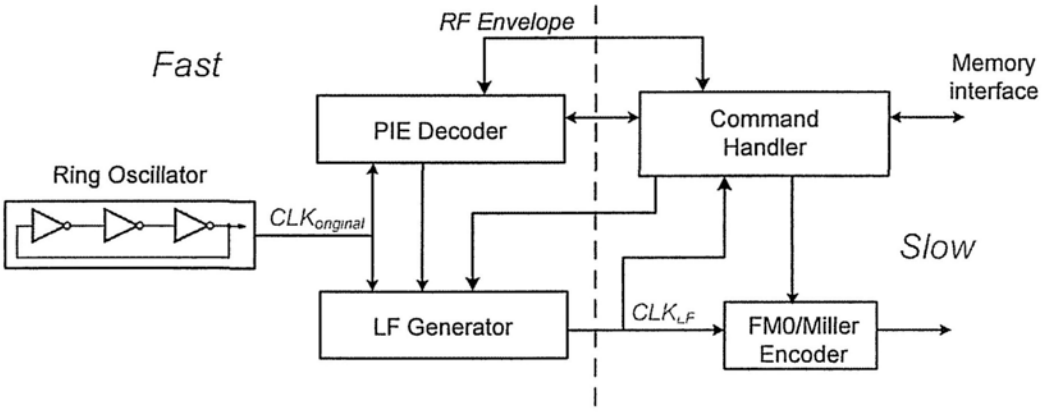


Figure 3.10 Initial block diagram in two clock domains

In fact the *Command Handler* processes and handles issues. It has to identify and verify the input commands, store and transfer the received timing reference from the *PIE Decoder*, access to the memory and response to valid command. The first two jobs can't accomplish without the synchronization of the edge of RF envelope. Also the *PIE Decoder* needs RF envelope to synchronize the symbol sampling. So in the advanced block diagram, the *RF Envelope* should be considered as a new clock domain.

As the whole BBP works in subthreshold region, it is better to pipeline the command handling to avoid calculation or processing congestion after the command is received. Otherwise the logic critical path will be too long in subthreshold operation. That means, the BBP can identify and process part of the already-received parameters in the payload, while the rest of the command is inputting in serially. Especially in receiving the *Select* command, the BBP may be required to compare up to 256 bit data between payload and the information from memory. Pipelining also shows more efficient in redundancy checks. For the random number generation, assigning the fastest RO-original clock to it can somehow provide more randomness.

With the consideration on all mentioned above, Figure 3.11 gives the optimized and detailed block diagram.

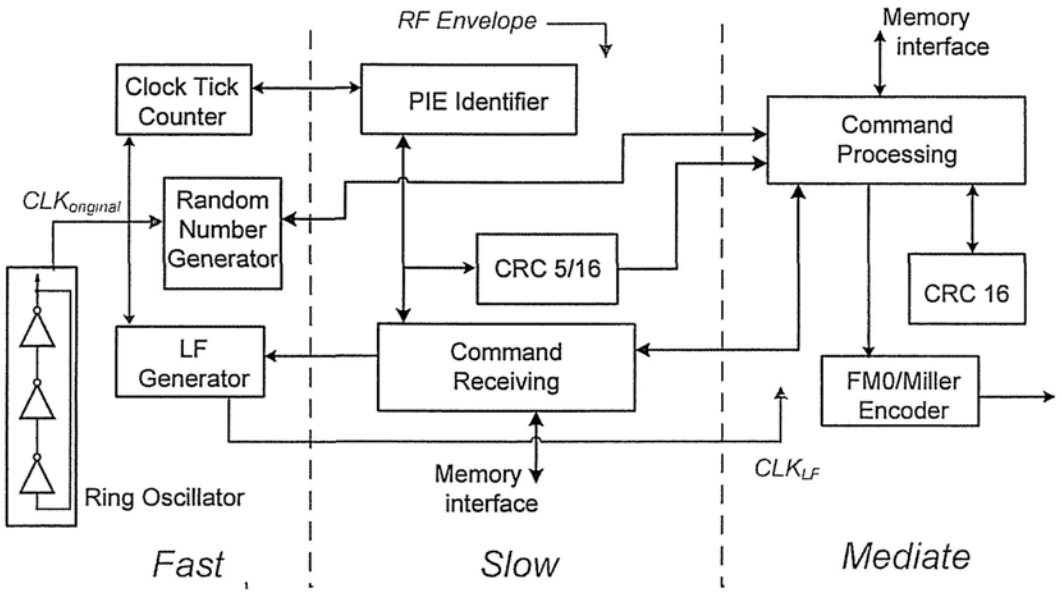


Figure 3.11 Optimized block diagram in three clock domains (for the purpose of clarity, the clock signals connections are not shown, denoting them in colored labels and arrows)

In Figure 3.11, *RF Envelope* is chosen as a clock signal, separating the original *PIE Decoder* and *Command Handler* into two parts. Now the *PIE Identifier* only focuses on counted value sampling, data decoding and header detection. The *Command Receiving* is responsible for the command type identification, parameters pre-processing, *LF* calculation and redundancy check control. And the *Command Processing* mainly takes charge of state transition, memory access and backscatter control. In this way the processing efficiency is increased. Adding the new slow clock somehow flatten the process distribution, it also liberate some components not necessary in high clock domains, reducing the risk of subthreshold timing violations.

A successful subthreshold BBP design is far more than merely block optimization in Figure 3.11. To ensure the design functions at very low supply, more elaborations

in basic levels are required. In the next two chapters, the related studies and novel designs are given in particulars.

CHAPTER 4 DESIGN ON LOGIC STRUCTURE FOR SUBTHRESHOLD & LOW POWER

4.1 Energy-aware and Pipelined Structure

Passive RFID tag is powered by no battery. There is no constant power source that is available in the near field. In another word, the energy for the baseband processor is directly dependent on the charge amount that is stored in the large capacitor at the LDO output, while the LDO relies on the rectifier's output power, and the RF power from the incoming wireless signal is the ultimately the source of power.

The most risky period in the subthreshold operation of BBP is in the reader-to-tag signaling. As the RF signal is modulated by PIE symbols, the amplitude of it has very low value when transmitting the low level part of data symbols (with a fixed interval *PW*). In this period the tag endures low availability of input power, and the instability of power supply is increased. Especially when a certain high energy/power is consumed in this moment, V_{DD} (the power supply) will sink fast to a level that is out of tolerance. Consequently it will cause logic errors.

The other risky period is in the backscattering period. The general mechanism of backscattering is by changing the input impedence look from the antenna. While the look-into impedence changed, in UHF domain, some power on the antenna will be reflected back. According to the strength of reflected-back power, the reader can identify the backscattered envelope is high or low. As some absorbed power is reflected out, when the large MOS in Figure 4.1 is ON, the power availability is low to the tag [40][41].

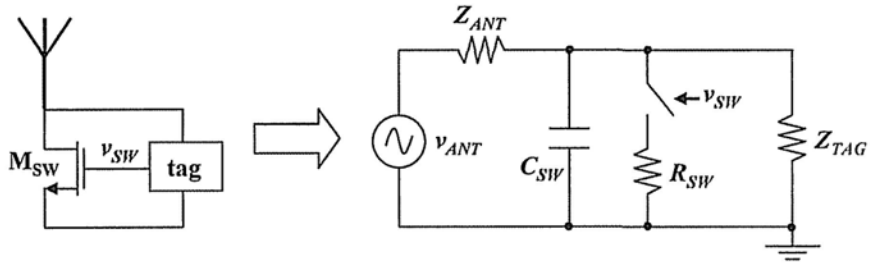


Figure 4.1 Modeling of the impedance of antenna

4.1.1 The Energy-aware *Command Receiving* Block & *RF Envelope* domain

Considering the ultra low speed and potential instability brought by the subthreshold power supply, the subthreshold BBP should try to arrange most of its logic activities in the period when harvested RF power is abundant. That means to consume energy at the high level part of PIE symbols when receiving command, the low level part of FM0/Miller encoded symbols when backscattering, and in the idle/waiting period when continuous full-amplitude RF wave (CW) is feeding the tag.

Review Figure 3.11, the slow clock *RF Envelope* is a clock signal with irregular period and duty, representing the incoming PIE symbols, therefore in the logic design of this clock domain, all the logic switching activities should be within the high value interval of *RF Envelope*. Without noise interference or environment change, gate circuit's logic switching is caused by its inputs. As the whole baseband processor, including the blocks fed by *RF Envelope*, is mainly a synchronize system, and in the purpose to avoid low power availability, all the sequential cells in the *Command Receiving* block are designed to triggered at the rising edge. Sequential cells' output are at the start point of logical paths, logic switching's propagation begins from here. If the high level interval of the *RF Envelope* is long enough, most propagation-induced logic activities can finish within this span. According to the protocol, the minimum interval of the high level in PIE symbol is 2.97 μs, and the general value is 4.25 μs. Even in subthreshold operation, gate delay is of 100 ns magnitude, the entire

propagation can finish before or just a bit exceeded to the falling edge of *RF Envelope*, by modifying the datapath. The *PIE Identifier* is also mainly triggered at the rising edge of *RF Envelope*, the detail design will be discussed in the next sub-chapter with the whole PIE decoding circuit.

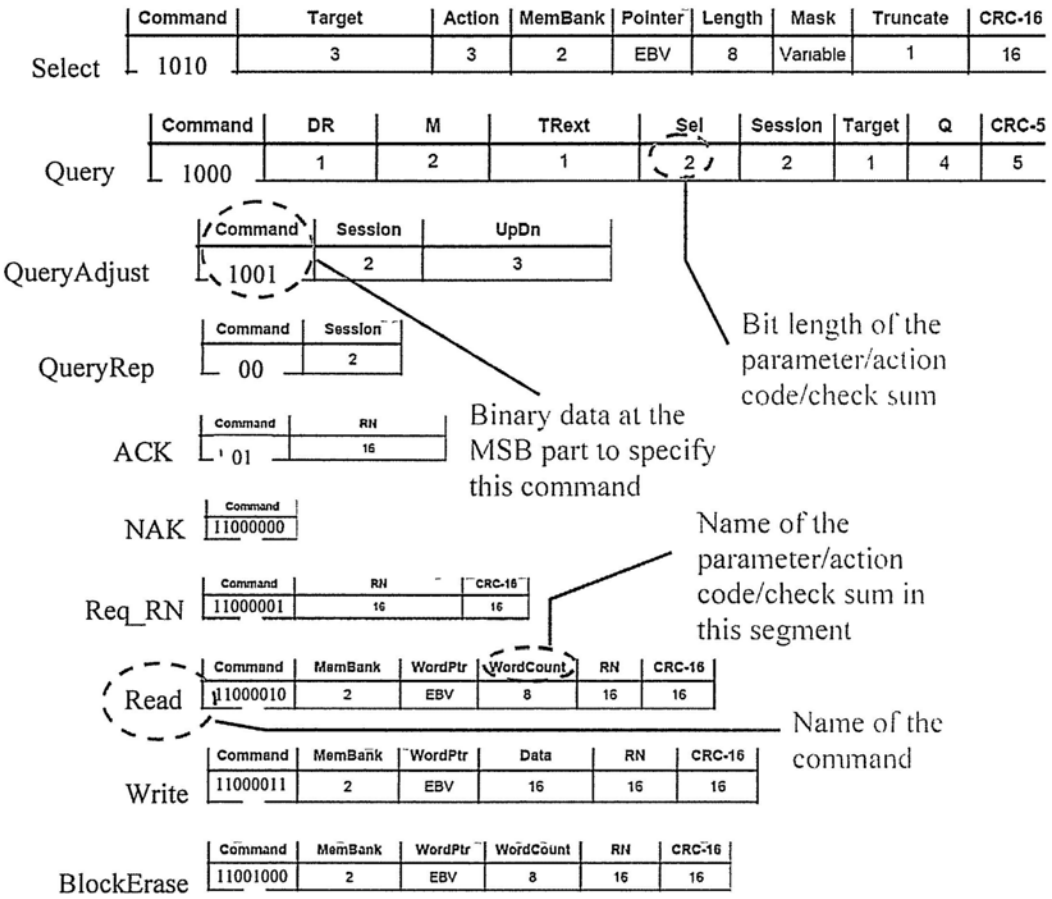


Figure 4.2 Main commands for the BBP

The function of the *Command Receiving* is firstly to identify the type of the command while shifting in the PIE decoded data, secondly is to pre-process and store the already received parameters. According to the protocol, commands are sent to tag with header at the first, and then the MSB of the payload (command). The content of commands is shown in Figure 4.2. The MSB 2~8 bit is used to define the specification of the command, followed by the related parameters. And most parameters, one or two or three together, can be divided into portions with 8-bit byte

length. Based on this pattern, the sequential cells in the *Command Receiving* block are organized in units of 8 bits. A binary counter is also needed to count the number of bits (symbol) that are already received. The state machine design is constructed dependently on the value of the counter. The values are in fact the states. If most of the values of the counter are designed as states, which are with data and status assignment to registers, the complexity could be high and the structure can be irregular. As mentioned above, since the registers are in 8-bit units, the states should also be based on the value of 8's common multiple. In another word, when the values of the counter is 8's multiple, some enabled 8-bit registers are assigned new data at the rising edge of *RF Envelope*, as well as some selected status/flag registers.

In subthreshold operation, logic delays are ultra long especially with high fan-out. Additionally, when the tag is comparatively far away from reader, the power availability can be quite low even when full-amplitude RF signal is at presence. Given that there are relatively too many registers are assigned new data at one of the *RF Envelope* rising edges when the counter is 8's multiple, the instant power consumption can be quite high after *RF Envelope* rises, timing violation is also with risk. Therefore in this work, the datapath and control design of the *Command Receiving* is not strictly based on 8's multiple. As shown in Figure 4.3, the register unit *shift_reg* is a 10-bit register that directly get the decoded binary data serially from the PIE decoder. With the size of 10 bit, the received data can now be transferred to the *byte_register 1*, *2*, *3* or *4* or the 6-bit discrete *parameter_register* at the rising edge of *RF Envelope* when the *symbol_counter*'s value (states) is 8's multiple, ±2 of the multiple, or some certain number not over 8 (because some commands' full length is less than 8 bit). Controlled in this pattern, the logic

switching is scattered to more triggered edges and the energy consumption is somehow flattened.

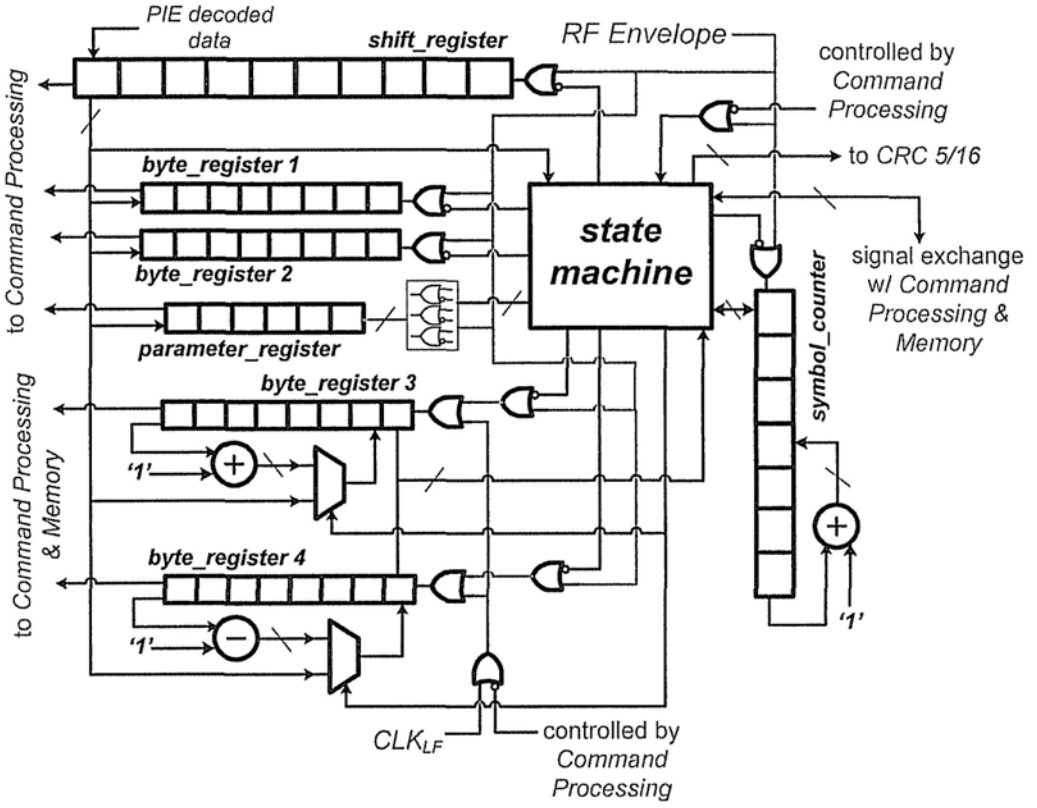


Figure 4.3 Power-saving structure of the *Command Receiving*

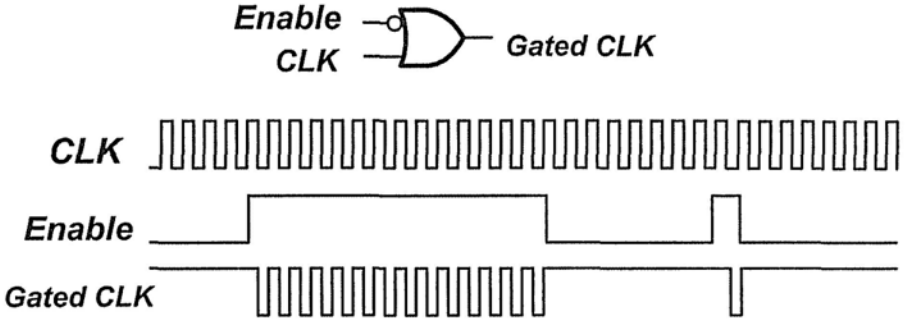


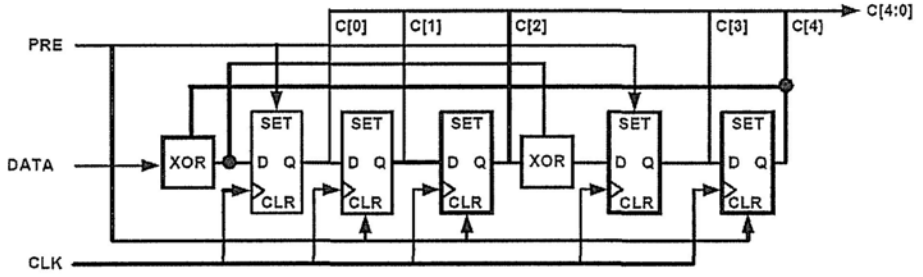
Figure 4.4 Clock gating behavior in *Command Receiving*

Battery-powered system design may targets on energy saving, regardless of instant power, but due to the specialty of passive RFID system, especially when modulated PIE symbols are being transmitted, power saving is the main target. Besides the flattening of power distribution, clock gating technique is also applied to control the

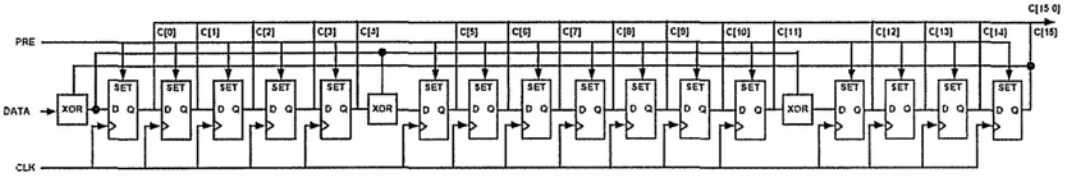
clock feeding to *byte_register 1*, *2*, *3* and *4*. As shown in Figure 4.4, the clock gate circuit controls the clock propagation that the gated clock has the fall-rise motion in every 8 or 10 "periods (symbols)" of *RF Envelope*. This structure avoids unnecessary charging/discharging at the clock nodes of the registers, and keeping the clock at high in idle time helps reduce dynamic power of the registers composed of rising-edge-triggered D flip flops.

As the "irregular" clock *RF Envelope* is a very low frequency signal, the subthreshold timing of the Command Receiving block is tense but not very critical. Data transferring among or inside *shift_reg* and *byte_register 1* and *2* is basically direct Q output to D input. The *byte_register 3* and *4* are a little bit complex with multiplexer and an 8-bit binary counters on the logical path between Q output and D input. The counters are used for memory access with variable bit length when the BBP is receiving *select* command or backscattering information for the received *read* command. One counter is with increment of 1 making *byte_register 3* as the address pointer, and the other one is with decrement of 1 to help *byte_register 4* to count down the length of bits that needed to access. At receiving the *select* command, the two registers will be firstly assigned the address and length, then they change the address and length with increment/decrement to access the memory, in this way the data read from memory can be compared with the data in select command bit by bit. Refer to the *read* command, the first step is the same, but the memory access is in the period of backscattering. Therefore the two registers work in two clock domain: *RF Envelo*pe and the CLK_{LF}, as the communication is half duplex, the two clock will not activated in the same time, so this arrangement is possible with not only lower power and complexity, also a bit help to release the timing tension in the *Command Processing* block. About the *symbol_counter* and the *state_machine*, in most of the

values of the counter, the *state_machine* does not give orders to registers for data assignment. Only in 8's multiple, ±2 of the multiple, or some certain number not over 8, as mentioned above, the *state_machine* sends new orders according to its current condition, the comparison results and received parameters. The *symbol_counter* is 7-bit with increment of 1 each time. In reader-to-tag signaling, the value of the counter will not proceed 64, and the values that are interested by the *state_machine* (the total states of it) is less than 25. In this structure, as shown in Figure 4.7, the critical path starts from the value output of *symbol_counter*, through state identifier (whether the value is related to a specified state), case matcher (to see whether all the state/parameter comparisons are satisfied), clock gating block, and finally the target register can be triggered and capture the data. The subthreshold (0.4~0.2 V) propagation time of the critical path is in 0.9~6.3 μs. As the minimum interval of the symbols in *RF Envelope* is 6.25 μs and the minimum power-abundant span is 3 μs, it can cover almost of the propagation time of the paths or motion described above. In the clock domain of *RF Envelope*, the power control is the primary issue while the subthreshold timing is the secondary importance.



(a)



(b)

Figure 4.5 Structure of (a) CRC 5; (b) CRC 16

The *CRC 5/16* redundancy checker are shown in Figure 4.5 (a) and (b) respectively. Reviewed, Figure 4.2, it is obvious that when the MSB 3 bit have been received, the baseband processor is able to decide whether activate CRC check and which CRC check should be selected. Therefore clock gating can help saving power.

4.1.2 The One-hot *Command Peocessing* Block & CLK_{LF} domain

In the link frequency domain (CLK_{LF}), subthreshold timing becomes high risky. And during the backscattering period, the harvested power will be partially reflected back when the original impedance is changed. In the command receiving period, the *Command Receiving* has already pre-processed the command identification, parameter verification and preparation. The *Command Processing* block is mainly for the tag state transition based on the command received and parameters provided, parameters assignment (especially the ones can label the uniqueness of the tag), and the arrangement of the data for backscattering. It also controls the *FM0/Miller Encoder* to backscatter data in correct format.

The protocol specifies the tag should reply (backscatter) to the reader after the reader-to-tag signaling is finished in 6.5 μs at minimum, while the LF is at the maximum 640 kHz (6.5 μs is from $MAX(RTcal, 10 \cdot LF^{-1}) \times (1 - FT) - 2\mu s$, FT is frequency tolerance, here is 0.15 for 15%). In subthreshold operation, the timing tension is very high. Fortunately the *Command Receiving* block has already pre-processed some data in advance. The *Command Processing* block starts to be clocked once the reader-to-tag talk is done. To finish the job of reply and leave some time margin for *FM0/Miller Encoder*, the Command Processing block should have the state transition and parameter assignment done within 5 periods of CLK_{LF}, and access the memory to give data to *FM0/Miller Encoder* thereafter. There is also a preamble and possibly a pilot tone before the MSB of the payload is backscattered (tag-to-

reader signaling is also a MSB-go-first format). And similarly there is a counter that the state machine control the whole block according to its current value. From being enabled, to finish backscattering, there will be 40~ 280 steps, actually the clock periods. Only the execution of *Read* or *ACK* command requires over-64 steps, because the BBP has to read the memory bit by bit with a large length. So the counter can reduce its counting range to 40~61 steps, by pausing the counting until the *byte_register 3* and *4* finish the memory access in response to *Read/ACK*. If binary counter is adopted, it should be a 6-bit counter. In a frequency of 640 kHz, the critical path starts from the output of the step counter, through state identifier, case matcher, data selector (multiplexer) and finally reaches at the D input of the register, of which Q output is connected to the *FM0/Miller Encoder*. In a 640 kHz clocking, the subthreshold delay of such critical path is likely to violate the timing requirement. Therefore the step counter of this work is embedded into the state machine, partially using the modified one-hot state machine structure.

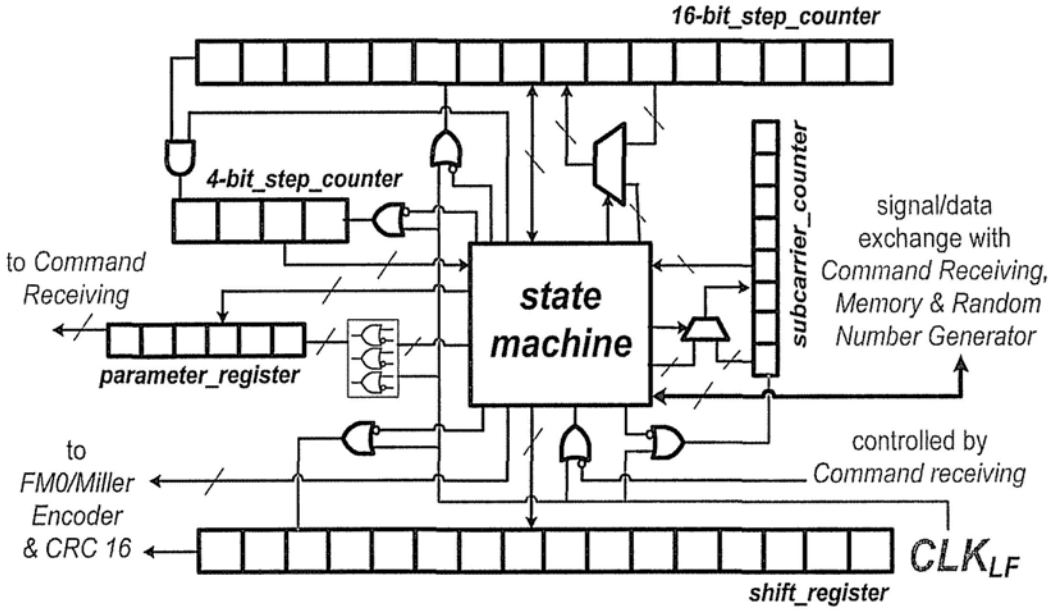


Figure 4.6 Structure of the *Command Processing*

As shown in Figure 4.6, the *16-bit/4-bit_step_counter* is in fact a state-indicator with two one-hot registers which are 4-bit and 16-bit respectively. The 4-bit one indicates the stage of processing, and the 16-bit one indicates the progress in a certain stage. One-hot structure state machine brings two benefits for subthreshold operation. First one is the step counter and the state identifier are somehow combined together with less complex logic function to identify the state, that means the great reduction of delay time in this part. The second one is the one-hot counter is basically a shift register that shifts in one direction, the delay path is a 2-in-1 or 4-in-1 multiplexer, which is shorter than an 6-bit counter. Figure 4.7 gives the demonstration. There is also an 8-bit one-hot counter *subcarrier_counter* for backscattering in Miller encoding, because Miller encoding uses *LF* as sub-carrier, A Miller sequence can contain 2, 4 or 8 sub-carrier cycles/bit. The merit of using one-hot counter here is similar. And the *subcarrier_counter* controls the shifting in the *step_counter* when backscattering Miller code.

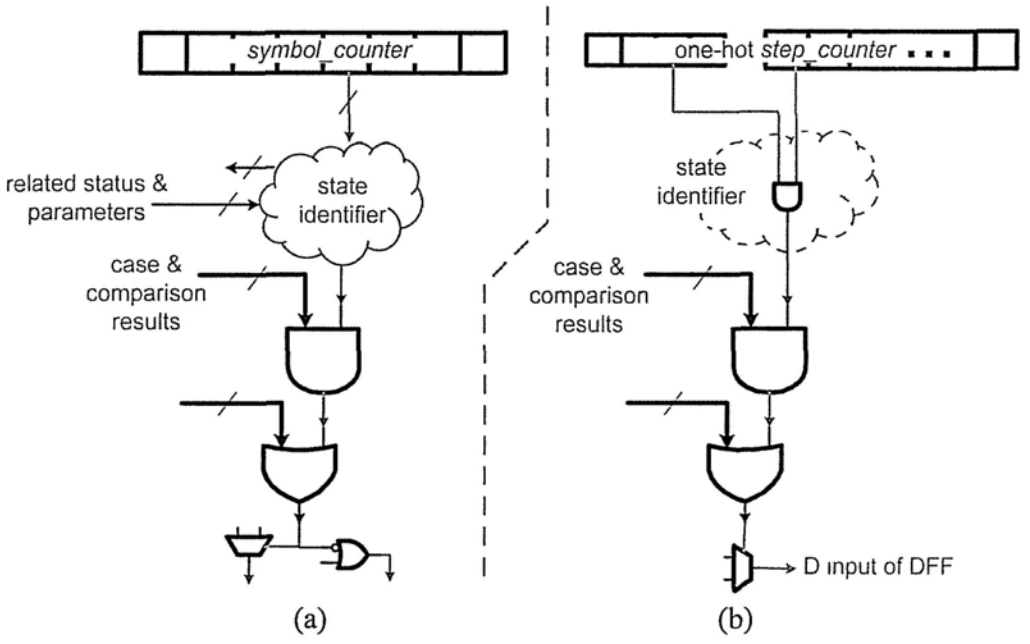


Figure 4.7 Critical paths of (a)the *Command Receiveing* (b)the *Command Processing*

The description of FM0 and Miller encoding is in Appendix D. The state transition on generating data symbols is not complicated and with the help of one-hot *subcarrier_counter*, the *FM0/Miller* Encoder is not risky in subthreshold operation.

4.2 Data Link Portion for Subthreshold Operation

The clock from the Ring Oscillator is the fastest clock in BBP. It is firstly utilized to count the length of the receiving PIE symbols, to help the PIE decoding block to calculate the timing reference *Tari*, *RTCal*, *TRCal* at the beginning, and calculate the data symbol's length to compare with *RTCal/2*, the *pivot*, and finally decide the binary value of the incoming PIE symbols. The clock is also used to feed the counter with increment of 1 each time, to compare with the *LF* period which is derived from the already-calculated *TRCal*, the clock CLK_{LF} for backscattering link frequency is therefore generated based on the comparison result.

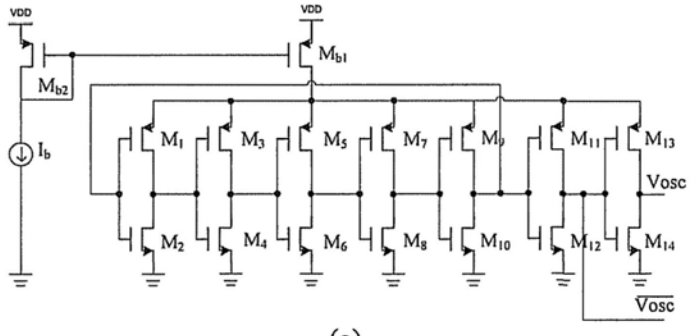
In subthreshold operation, it would be easier if the fastest clock goes slower. However, according to the protocol, the $CLK_{original}$, which is from the *Ring Oscillator*, must keep above a certain frequency. To achieve good PIE decoding on data symbols with low bit error rate (BER), high accuracy with high time resolution is required. The minimum clock frequency to drive a binary counter for PIE decoding must be more than 1.6 MHz [44]. Considering the data rate accuracy for uplink, the backscattering using CLK_{LF} for link frequency, generally the design must select a minimum clock frequency above 1.98 MHz, and 2.56 MHz is a preferred choice offering better margin [45].

4.2.1 *PIE Decoder* for Subthreshold Timing

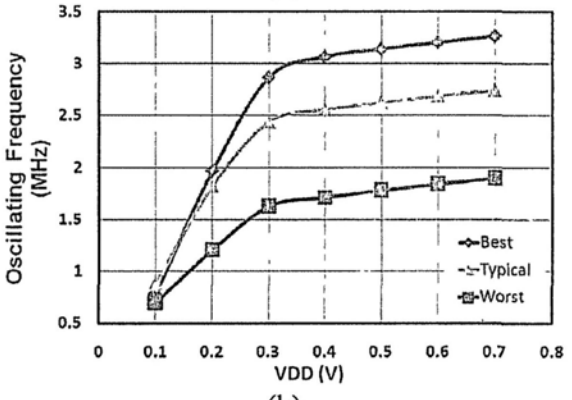
1.98 MHz is a piece of cake in super-threshold operation, but a rough stone in subthreshold. At 0.3 V power supply, merely a general 2-input gate delay can be over

300 ns, without an elaborative design, the logic propagation can't accomplish within 500 ns. And the frequency is not guaranteed at the same very low power supply.

To save the power and chip area, the Ring Oscillator adopts a simple and general current starved structure [46], as shown in Figure 4.8(a). The RO is common and easy to design, however, it is another tough stone for PIE decoding. Due to PVT variation, it is still very difficult to ensure its oscillating frequency can be kept in an acceptable range when the center frequency and the power supply are low in values for low power purpose. The *Ring Oscillator (RO)* was already designed in the target technology. Figure 4.8(b) shows the curves of oscillating frequency of the ring oscillator in different cases, the *RO* is set to center at 2.56 MHz and 0.4V supply in typical case. This is a post-layout simulation and it is obvious that in the worse case, the frequencies are all below 1.98 MHz when the supply voltage is less than 0.7 V, and the maximum variation can be as high as 53%.



(a)



(b)

Figure 4.8 (a) the *structure of Ring Oscillator;* (b) The output frequency of *RO* in cases

To cover the entire protocol, the counter in PIE decoding needs to calculate a maximum symbol time interval of 225 μs. It occurs when *Tari = 25 μs*, *RTCal = 3Tari*, *TRCal = 3RTCal*, *TRCal = 225 μs*. And the protocol specifies the interval length of 4RTCal as an error received symbol. So it is required to design a counter capable of counting in 6.25 μs (the minimum *Tari*) ~ 300 μs. Given the *CLKoriginal* can keep as low as 2 MHz, it still need a 10-bit binary counter to cover the digitized amount, plus the instability of Ring Oscillator, the 0.2~0.4 V subthreshold operation of this PIE counter is mission impossible.

Therefore in this work, to maintain a low BER in the worst case and the voltage supply can be kept low, a mixed counter is proposed for the PIE decoder. As shown in Figure 4.9(a), the counter in PIE decoder has a mixed binary-ripple structure. The least significant 2 bit of the counter is composed of 2 DET (double-edge-triggered) flip flops and an adder that adds one to the two output bits taking as a binary number. In this way the clock signal COUNTER_CLK from the ring oscillator can trigger these two flip flops at both rising and falling edges, and the outputs of #1 and #0 flip flop will give a cyclic "00-01-10-11" for every two clock periods. The other part of the mixed counter is formed by a 9-bit ripple counter with single edge triggered (SET) flip flops, and the LSB of it (#2 flip flop, SET) is clocked by the output of #1 flip flop (DET). The mixed counter is functionally equivalent to a SET 11-bit ripple counter or 11-bit binary counter, but it halves the clock frequency for the same time precision.

The main advantage of this mixed counter is the reduced variation with a lower frequency. Compared to Figure 4.8(b), the ring oscillator is now equivalently set to center at 1.28 MHz and 0.4V supply in typical case for the proposed mixed counter. Figure 4.9(b) shows the post-layout curves of frequency variation in different cases.

It is shown that when the ring oscillator centers at 1.28 MHz, the oscillating frequency in the worst case does not drop too low to keep above 0.99 MHz, which is equivalent to 1.98 MHz for SET binary or ripple counter. And the case-to-case frequency variations are now less than 22%. In addition to this improvement, low clock frequency can also reduce power consumption for both the ring oscillator and the *Mixed Counter*. More importantly, by cooperating with the rest part of the PIE decoder in a certain manner, the subthreshold timing problem can also be solved.

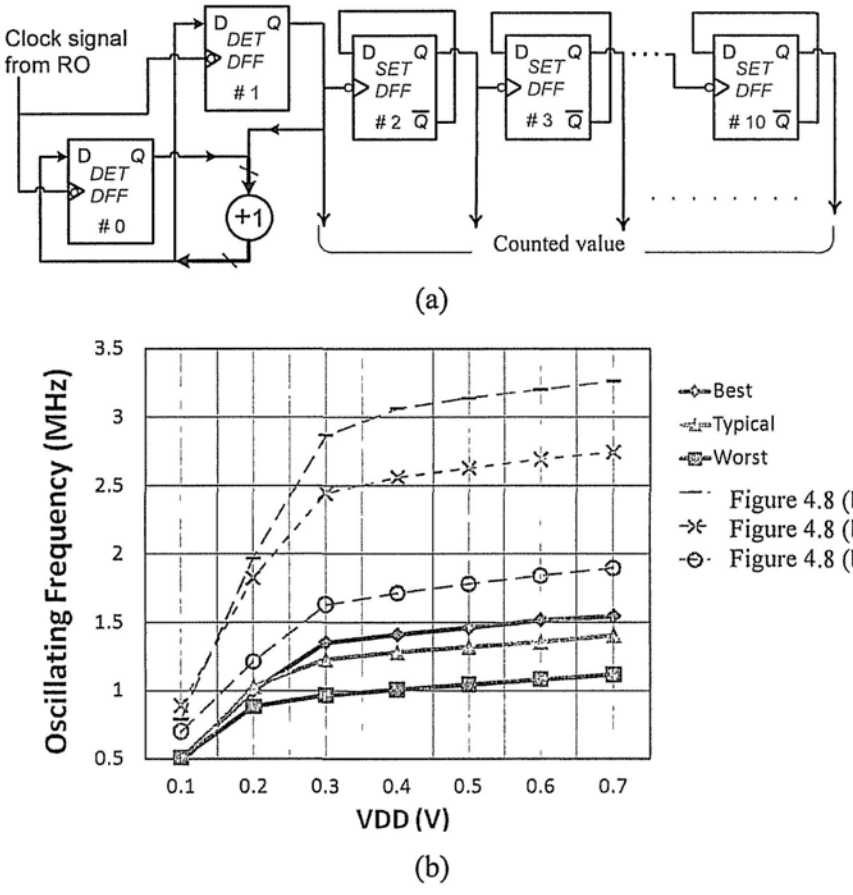


Figure 4.9 (a) the *structure of Mixed Counter;* (b) The output frequency of *RO* in PVT variation, centered at 1.28 MHz, 0.4V power supply, typical case

Like the *Command Receiving* block, the rest part of the PIE decoder features an energy-aware structure. As shown in Figure 4.10, the irregular clock *RF Envelope* is in fact the extracted envelop of RF signal. Energy awareness is achieved by arranging

most of the data transfers and calculations with the irregular clock. All the registers controlled by *RF Envelope* are triggered at the rising edge; the moment when the harvested RF energy begins to be abundant. As the phase of the $CLK_{original}$ is not predictable and controllable, the *Gating Clock* module here uses D flip flop to avoid glitch and racing condition. Figure 4.11 shows its structure and behavior.

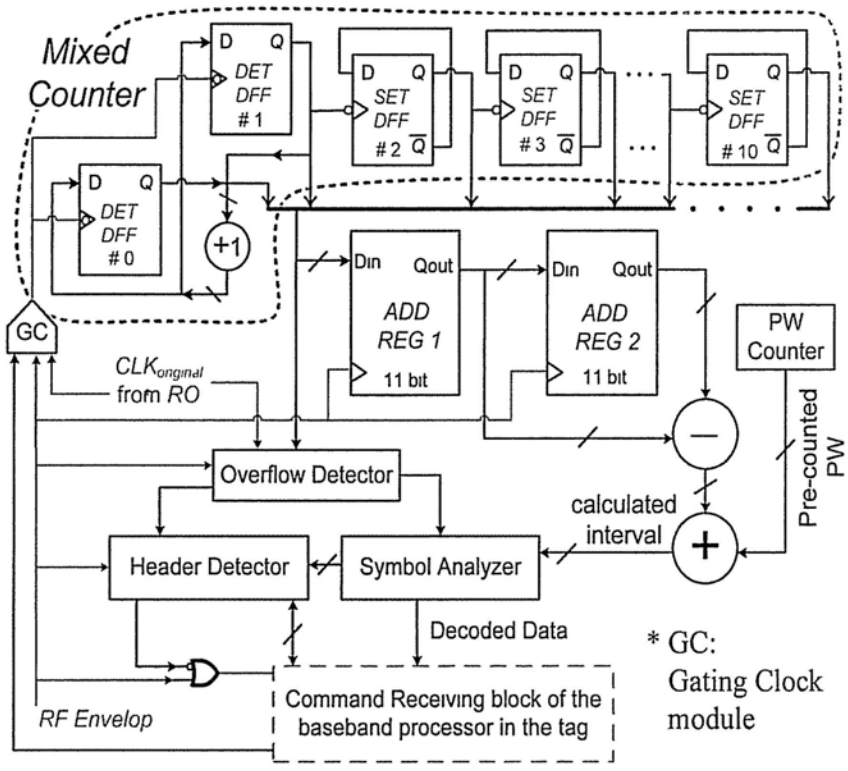


Figure 4.10 Energy-save and pipeline structure of PIE Decoder

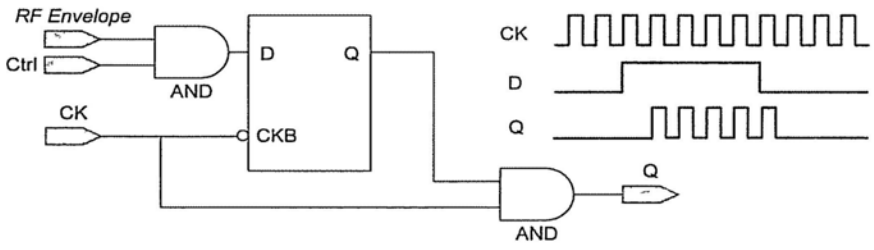


Figure 4.11 Clock Gating module for PIE decoder

According to the C1G2 protocol, *PW* in Figure 3.3 has a constant interval for every communication round [43]. So in this design, the *Mixed Counter* only counts in the interval of every symbol when *RF Envelope* is at high level; *PW* is pre-counted

by another 5-bit ripple counter at the first symbol of an incoming command, and then input to the adder-subtractor unit as a temporarily constant value (the *Pre-counted PW* shown in Figure 4.10). The whole design releases the timing constraint on the mixed counter at low-voltage supply with fast clock input: ripple propagation in Figure 4.9(a)'s sub ripple counter is nothing to do with the *CLKoriginal*, and if the state propagation from #2 to #10 is required when *RF Envelope* goes down, it can also be done when the *RF Envelope* is low (because *PW* is at least 2 μs long, enough even in subthreshold region), critical path in *CLKoriginal* domain is reduced to a 2-bit half adder's carry-to-sum delay or D flip flop's CK-to-Q delay, when compared to a 11-bit binary counter. Figure 4.12 demonstrates the pipelined PIE decoding process of the decoder.

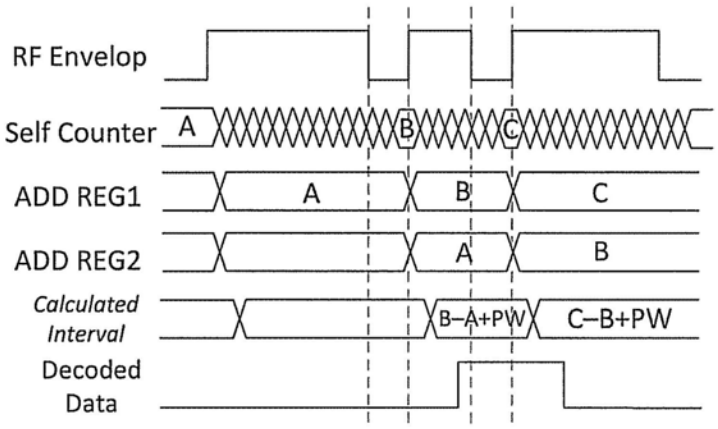


Figure 4.12 Pipelined and energy-aware process of PIE decoding

Utilizing the double edge triggered counting, the *PW* pre-counting and the gating clock method can release the subthreshold timing and achieve energy saving. But as only the high value part is measured for every symbol, and the clock gating is adopted, numerical simulations are required to ensure that the decoder can correctly identify the data-0 and data-1 symbol. There is a key parameter called Decoding Margin for the evaluation of PIE decoding. It is a reference to the number of clock ticks that separate *RTcal/2* (*pivot*) from the measured symbol. For the ordinary binary

counter that calculates the whole interval of the symbol by single clock edge, the decoding margin at 1.98 MHz is shown in Figure 4.13 [45]. In the protocol-defined range of *Tari*, as the calculated values of data-0 (*Tari*) and data-1 (*1.5Tari*) symbols keep minimum 1 tick away from the pivot (above 0 in decoding margin), it indicates that the ordinary counter can correctly indentify them by digitized comparison.

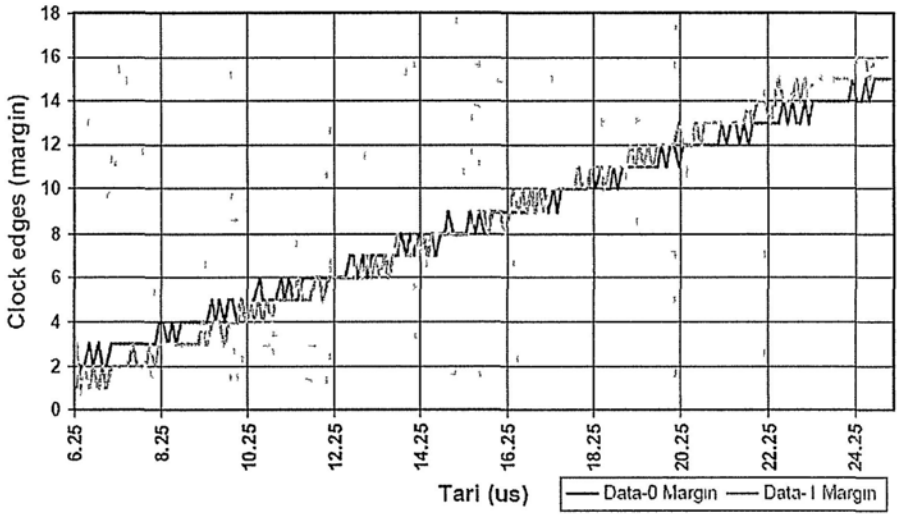


Figure 4.13 Decoding Margin versus the *Tari* at 1.98 MHz from [45]

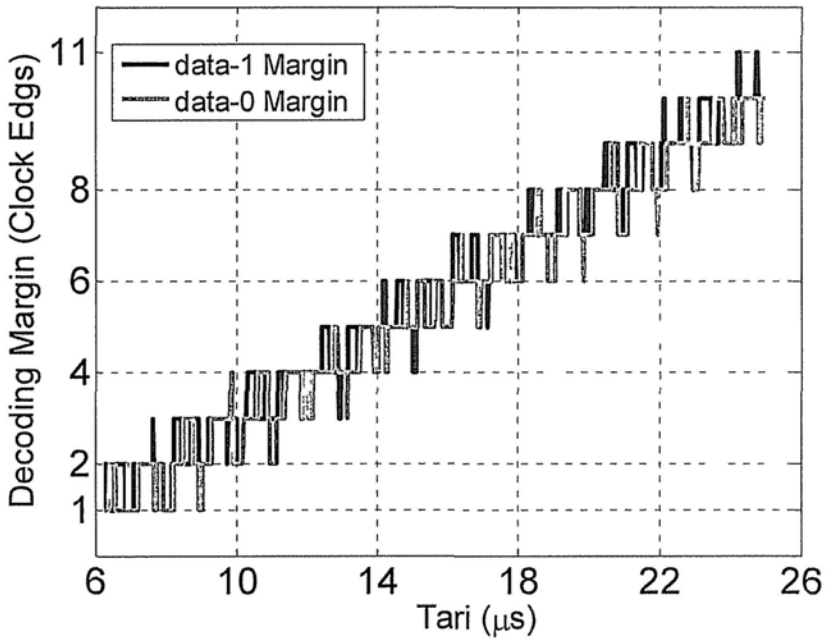


Figure 4.14 Worst-case Decoding Margin versus the *Tari* of the proposed PIE Decoder

According to the logic structure of this PIE Decoder, a detailed numerical analysis is carried out to calculate the decoding margin. Based on the frequency variation from 0.3 V to 0.7 V in Figure 4.9(b), with 1000 selected points from the range of *Tari*, the minimum numerically-possible values of the decoding margin for data-0 and data-1 were plotted in two lines. Figure 4.14 is the simulations result, indicating that the proposed PIE decoder can correctly function.

4.2.2 *LF Generator* for Sunthreshold Timing

Generally an LF generator uses clock ticks to count and controls the rising and falling of the CLK_{LF} according to the comparison between its counted number and the previously-stored number LF^{-1}, which is calculated and quantized from Equation (3.1). Subthreshold timing violation is in the clock tick counter that is fed by the *Ring Oscillator (RO)*. According to the protocol, the PIE timing reference *TRCal* is 200 µs at maximum. There are also two data rate which are specified in *DR* of the *Query*

Table 4.1 Tag-to-reader Link Frequency Requirment

DR: Divide Ratio	TRcal[1] (µs +/– 1%)	LF: Link Frequency (kHz)	Frequency Tolerance FT (nominal temp)	Frequency Tolerance FT (extended temp)	Frequency variation during backscatter
64/3	33.3	640	+/– 15%	+/– 15%	+/– 2.5%
	33.3 < TRcal < 66.7	320 < LF < 640	+/– 22%	+/– 22%	+/– 2.5%
	66.7	320	+/– 10%	+/– 15%	+/– 2.5%
	66.7 < TRcal < 83.3	256 < LF < 320	+/– 12%	+/– 15%	+/– 2.5%
	83.3	256	+/– 10%	+/– 10%	+/– 2.5%
	83.3 < TRcal ≤ 133.3	160 ≤ LF < 256	+/– 10%	+/– 12%	+/– 2.5%
	133.3 < TRcal ≤ 200	107 ≤ LF < 160	+/– 7%	+/– 7%	+/– 2.5%
	200 < TRcal ≤ 225	95 ≤ LF < 107	+/– 5%	+/– 5%	+/– 2.5%
8	17.2 ≤ TRcal < 25	320 < LF ≤ 465	+/– 19%	+/– 19%	+/– 2.5%
	25	320	+/– 10%	+/– 15%	+/– 2.5%
	25 < TRcal < 31.25	256 < LF < 320	+/– 12%	+/– 15%	+/– 2.5%
	31.25	256	+/– 10%	+/– 10%	+/– 2.5%
	31.25 < TRcal < 50	160 < LF < 256	+/– 10%	+/– 10%	+/– 2.5%
	50	160	+/– 7%	+/– 7%	+/– 2.5%
	50 < TRcal ≤ 75	107 ≤ LF < 160	+/– 7%	+/– 7%	+/– 2.5%
	75 < TRcal ≤ 200	40 ≤ LF < 107	+/– 4%	+/– 4%	+/– 2.5%

command, from Equation (3.1) and the protocol requirement, Table 4.1 gives the timing requirement of the link frequency [43]. The minimum *LF* defines the

maximum LF^{-1} 25μs, and given the maximum $CLK_{original}$ is 3.3 MHz, the binary counter for *LF* must be at least 6-bit long. Consider that the supply voltage is subthreshold low, carry propagation in the 6 bit binary counter can increase to hundreds of nanoseconds, counting error may occur if this delay is longer than the period of the relatively high-frequency input clock $CLK_{original}$.

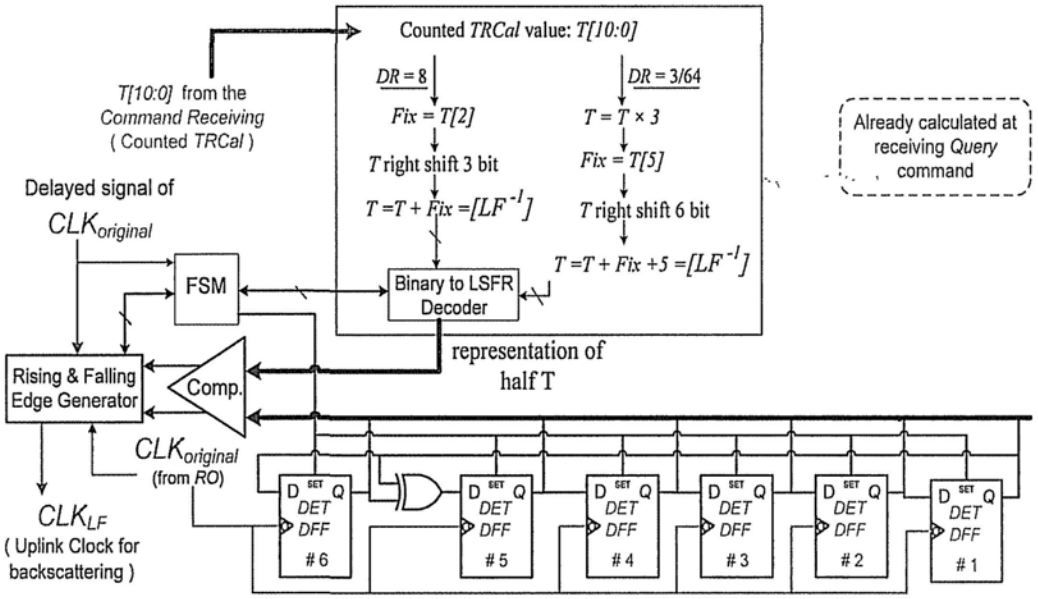


Figure 4.15 Structure of LF Ganerator

For the purpose of removing carry propagation, Galoi linear feedback shift register (LFSR) is utilized in the *LF Generator* as a counter. The merit of LSFR is that the long carry propagation delay is replaced by a simple XOR gate. Figure 4.15 shows the structure of the *LF Generator*. Here the polynomial for the Galoi maximal LFSR counter is X^6+X^5+1, and a binary-to-LFSR decoder is used on the non-critical path. In the *PIE Decoder*, *TRCal* is counted as a certain number of *CLKoriginal*'s half period, so this generator also uses clock-efficient DET flip flops as sequential cell. According to the EPC C1G2 protocol [43], *DR* can be 8 or 64/3, and the frequency-variation requirement of the *LF* for backscattering is critical for the low-voltage baseband processor, especially when *DR* is 64/3, *LF* is 640 kHz and the *Ring*

Oscillator offers a slow clock (*CLKoriginal*). A simple algorithm is proposed for the uplink clock generator to calculate *LF*'s period LF^{-1} into a rounded number of clock edges. As shown in Figure 4.15, a parameter called *Fix* is applied to reduce error of the result $[LF^{-1}]$, which is used for comparison. *Fix* is an important compensation as the error is mainly from the *TRCal*, induced by the missed or extra clock ticks in the counting, and the truncation in calculation.

By modeling the proposed *TRCal* counting and $[LF^{-1}]$ calculation with worst-case error and in the $CLK_{original}$ range from 0.5 MHz to 1.92 MHz, Figure 4.16(a) and (b) shows the selection of the clock-tick's number from the proposed *LF Generator*, as the frequency varies, and the target *LF*s are from the featured *LF*s in Table 4.1. It indicates that for both *DR* values, when the input clock is above 0.9 MHz, this simple design can give a correct clock tick number selection $[LF^{-1}]$ for every featured *LF* values in Table 4.1, with acceptable frequency tolerances specified in the protocol. Equivalently it has a smaller minimum clock frequency, 1.8 MHz, than what is suggested in [45], 1.98 MHz.

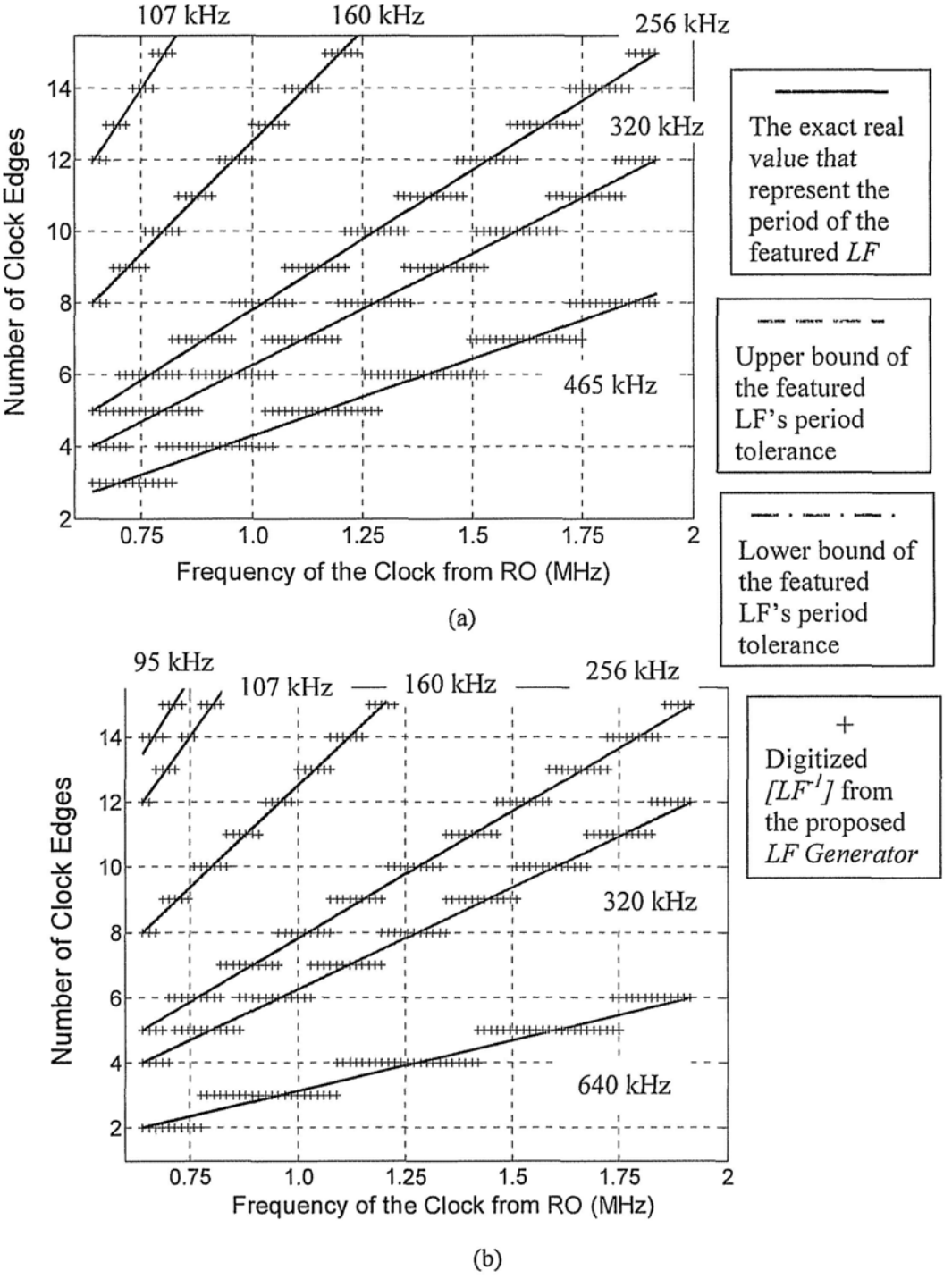


Figure 4.16 Calculated *[LF-1]* from *LF Generator* vs. frequency of the $CLK_{original}$ (a) *DR=8*; (b) *DR=64/3*

4.3 Entire Operation of Subthreshold Baseband Processor

With the purpose of power awareness and low power in subthreshold region, now the entire structure of the baseband processor is shown in Figure 4.17.

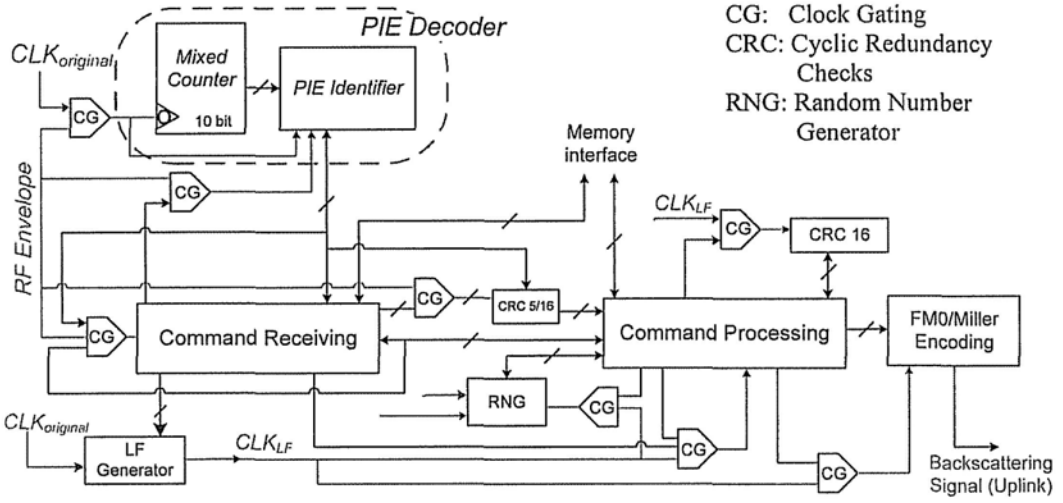


Figure 4.17 Structure for power-saving subthreshold Baseband Processor

The subthreshold power supply suppresses the static and dynamic current of the system, and the gating clock modules control the dynamic power for energy/power saving. The operation summary and flow of the whole BBP can be viewed in Table 4.2 listed below.

Table 4.2 Operation Step and Active Blocks

STEP / BLOCK	Idle	Raeder to Tag Comm.	Reader waiting Tag	Tag to Reader Comm.
Mixed Counter	ON	ON	OFF	OFF
PIE Identifier	ON	ON	OFF	OFF
CRC 5/16	OFF	ON *If required	OFF	OFF
Command Receiving	OFF	ON	OFF	OFF
LF Generator	OFF	OFF	ON	ON *If required
Command Processing	OFF	OFF	ON	ON *If required
CRC 16	OFF	OFF	OFF	ON *If required
FM0/Miller Encoder	OFF	OFF	OFF	ON
Random Number Generator	ON	OFF	ON *If required	OFF

CHAPTER 5 ACCUSTOMED SUBTHRESHOLD LOGIC CELLS

Logic structure optimization is an important way to release the critical timing in subthreshold operation. As the supply voltage keeps scaling down, the pipelining, parallelism or algorithm techniques can't efficiently reduce the exponentially-increased path delay. The transistor-level elaboration of the custom subthreshold cells is an alternative assistance for the digital design to "dive deeper" at power supply. Chapter 2 has already introduced some state-of-art logic circuit designs for subthreshold usage, in this chapter, based on the generality in digital circuits and the actual requirement from the example design of this work: the subthreshold baseband processor, an innovative logic style is proposed to achieve faster propagation at very-low power supply: the Active Controlled Ratioed Logic (ACRL). It is a combination of the Pseudo-NMOS logic and the source follower pull-up logic (SFPL), tailored for the circuits where multiple path selection, magnitude comparison or fast sequential cells are taking the important role in timing requirement. The following sub-chapters are going to describe the topology, mechanism and give examples of this style.

5.1 Basics of ACRL

5.1.1 Introduction of Ratioed Logic

Ratioed circuits use weak pull-up transistor and stronger pull-down transistors. They reduce the input capacitance by eliminating large PMOS transistors loading at the input terminals, and hence the logical effort of the logic cells is improved. But the functional correctness depends on the right ratio of pull-up to pull-down strength. If the pull-up is too strong, the maximum output of logical low-level V_{OLmax} will be too

high, and conversely the minimum output of logical high-level V_{OHmin} can be too low, additionally the rising delay is too slow.

In super-threshold operation, V_{OLmax} is required to be less than the threshold voltage so the low output does not turn ON the next stage logic gates. And the static power is inevitable when the output is low. Therefore in super-threshold usage ratioed circuits are adopted in a limited fashion where their benefits are significant. [47]

The most common form of MOS ratioed logic is the pseudo NMOS logic gates. The pull-down network is like that of a static CMOS gate, but the pull-up network has been replaced with a single PMOS transistor that is gate-grounded so it is always ON. The PMOS transistor width is selected to be about 1/4 the strength (i.e., 1/2 of the effective width) of the NMOS pull-down network as a compromise between noise margin and speed; this best size is highly process-dependent, but is usually in the range of 1/3 to 1/6 [47].

When power supply drops into subthreshold region, pseudo-NMOS may fail to function as weak robustness. In subthreshold region, the V_{OLmax} is not just required to be less than Vth, it should be less than 1/8 of the full voltage swing, otherwise it may violate the noise margin of the next stage. Very-large size of NMOS network can pull V_{OLmax} lower, but the penalty of area and rising delay time are more evident. Figure 5.1 shows the V_{OLmax} of an AOI (AND-OR inverter) pseudo-NMOS logic gate versus the effective width ratio of PMOS to the NMOS in pull-down network. From the figure it can be observed that the values of V_{OLmax} at different VDD keep above the 1/8 voltage swing lines in most of the strength (width) ratio, even the area and power limit can be extended, large size NMOS network causes poor rising delay and therefore the entire performance.

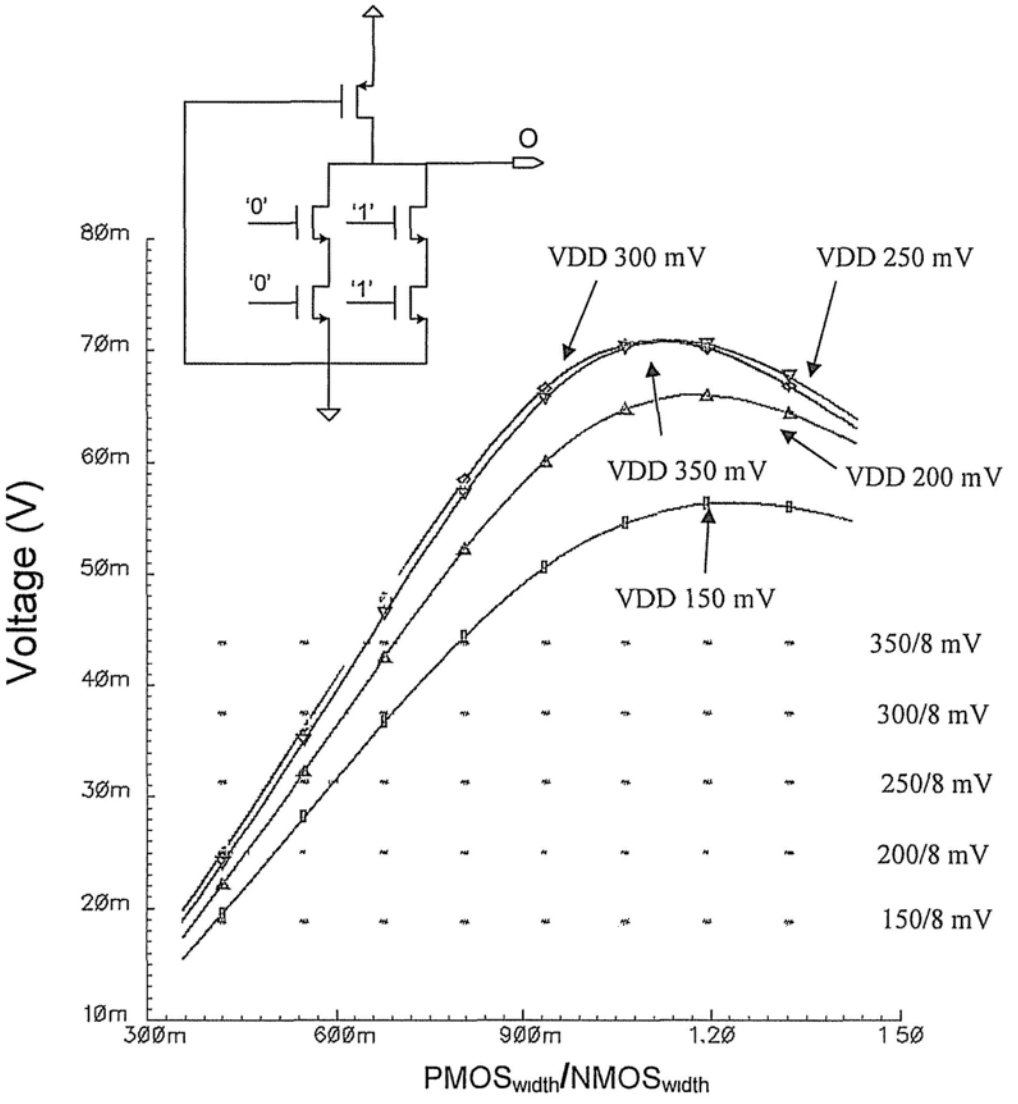


Figure 5.1 V_{OLmax} vs. effective PMOS/NMOS width ratio in different VDD of a general gate of pseudo-NMOS

Figure 5.2 shows a Source Follower Pull-up Logic (SFPL) 4-input NOR gate [48]. It is similar to a pseudo-NMOS gate except that the pull-up is controlled by the inputs. N6-N9 and P1 form a pseudo-NMOS NOR function. The gate of the pull-up P1 is driven by a parallel source follower consisting of drive transistors N1-N4 and load transistor N_{load}. When one input turns on, the source follower pulls node x to approximately VDD/2 [48]. This tends to partially turn off P1, which allows smaller NMOS pull-downs N6-N9 to be used. SFPL was firstly proposed in 1992, it was

merely used to constructing wide NOR gates, and not a general structure for super-threshold designs as the ultra high power consumption.

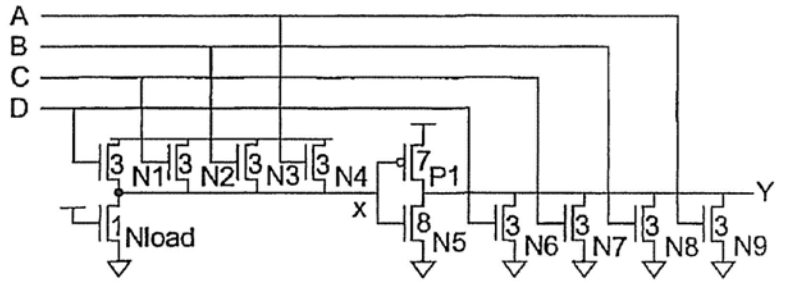


Figure 5.2 4-input NOR gate of SFPL style

5.1.2 Improvement of ACRL

Both of the pseudo-NMOS and SFPL has merits suitable for subthreshold operation. The important one is that the superiority of NMOS transistor. Figure 5.3 is the ON current comparison of PMOS and NMOS devices of target technology, derived from Figure 2.28~2.30. Even though the NMOS current is not 2~3X of the PMOS current in super-threshold region, NMOS is still a stronger device. In the

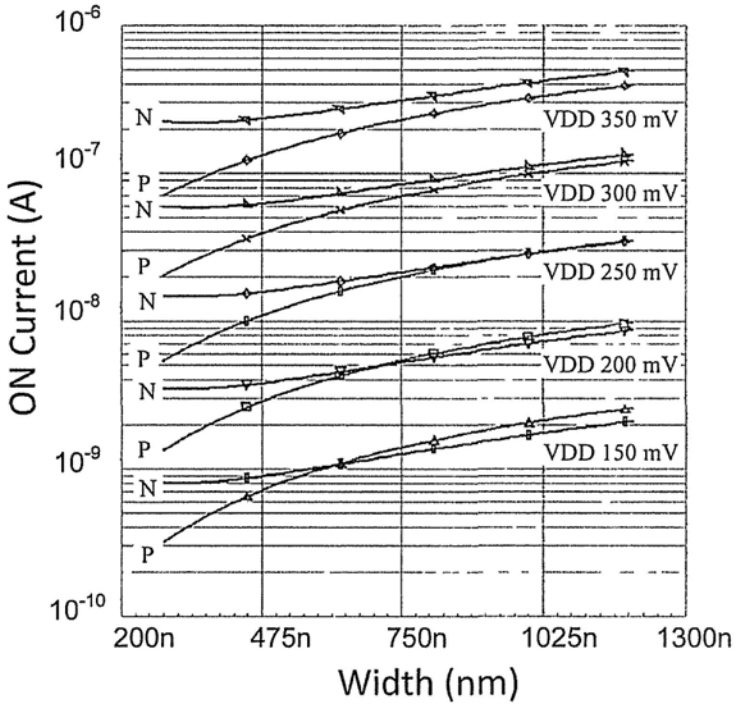
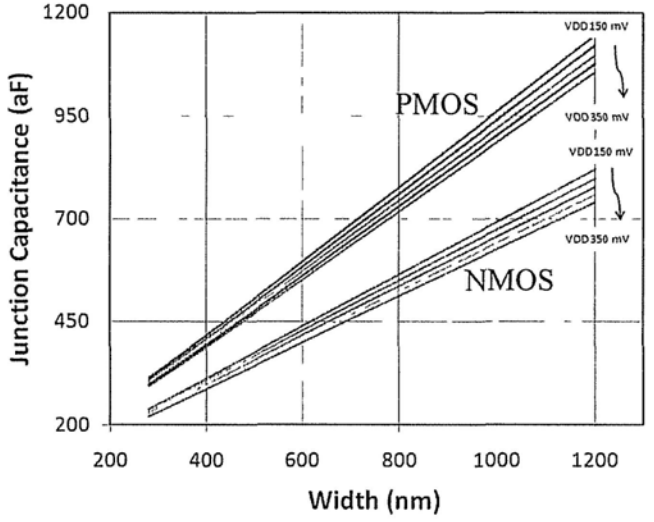
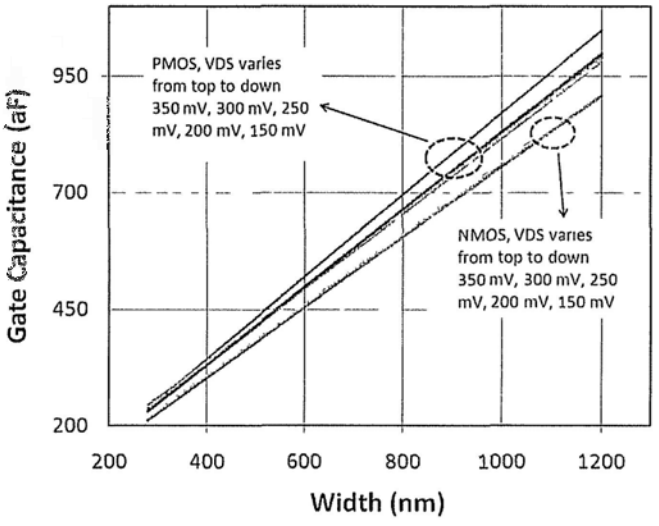


Figure 5.3 ON current of PMOS and NMOS transistors vs. width at different VDD

aspect of capacitance, NMOS also has smaller amount of the gate input capacitance and the junction capacitance. Figure 5.4 gives the capacitance comparison of the two devices in the target technology with different condition. Gate capacitance affects the



(a)



(b)

Figure 5.4 PMOS and NMOS capacitance comparison of (a) active area junction capacitance; (b) gate capacitance

speed of the voltage change of VGS and smaller junction capacitance can speed up the voltage change at nodes, plus the NMOS current superiority and the fast response of the active pull-up load, ratioed circuits have potential advantage for subthreshold design. With the combination and improved modification of the two structures, a

novel logic style Active Control Ratioed Logic (ACRL) is proposed in this work to advance the subthreshold designs. The improvement work is based on three issues:

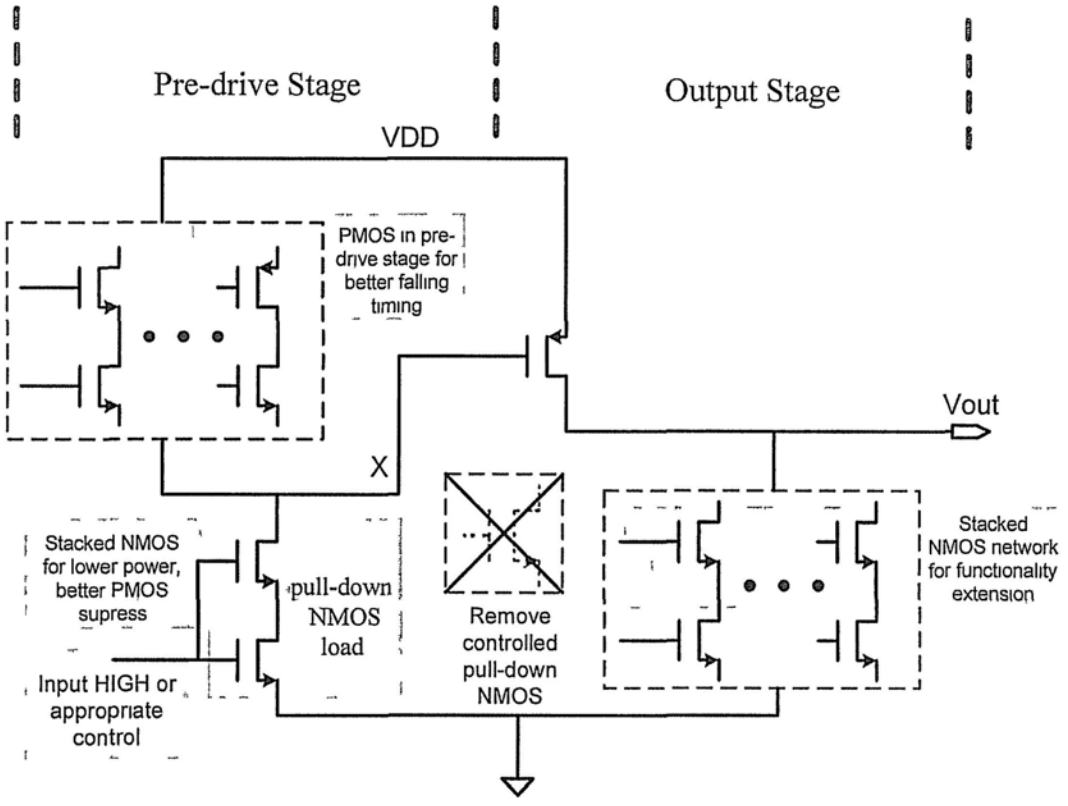


Figure 5.5 Improvement of proposed ACRL structure for subthreshold design

(1) The ACRL has the pre-drive stage similar to the source follower in SFPL, but the controlled pull-down NMOS in SFPL is removed and only the single pull-up PMOS is active, like the pseudo-NMOS structure. The reason of removing is for power reduction and timing improvement. In subthreshold operation, the node x, as shown in Figure 5.5, is not able to go up near VDD/2 when the path in pull-up NMOS network of the pre-drive stage is ON. As all the transistors work in subthreshold region, the pull-up strength of the NMOS network in pre-drive stage is exponentially sensitive to the VGS, even the node x just rises up to few dozen mV, the ON current of pull-up NMOS will degrade so much that the strength fighting at node x is balanced. In this circumstance, the current of

the controlled pull-up PMOS is several magnitude higher than the controlled pull-down NMOS, so this NMOS is removed with extra timing bonus as both the load capacitance at node x and output are reduced. Another optional improvement is that the original NMOS load N_{load} in Figure 5.2 is stacked with one more NMOS transistor for a higher balanced level in strength fighting. It is also a consideration for lower power and better falling timing.

(2) In the original SFPL style, only the NOR gate structure was proposed in application. That means the branch path from power supply to ground is two stacked NMOS or one PMOS and NMOS device. In the proposed ACRL structure, the NMOS network can be composed of two or three NMOS stacked branches. The relative pull-down strength will not be reduced as the node x can actively tune the current strength of the pull-up PMOS. As described in (1), when the pull-up network in pre-drive stage works, node x will rise up to approximately 20~35 mV, this slight change in subthreshold region causes 50~80% change of the PMOS current. Figure 5.6 shows the V_{OLmax} of an AOI (AND-OR inverter) ACRL logic gate versus the effective width ratio of PMOS to the NMOS in its pull-up and pull-down network. It has the same logic function of Figure 5.1 but apparently different results. The ACRL V_{OLmax} outputs are below VDD/8 in most of the condition. The availability of stacked transistor in input networks greatly extends the functionality of ACRL logic. Similar but more robust than pseudo-NMOS logic, the ACRL can maintain the performance with reliability.

(3) Some logic or sequential function need signal inversion to the input network. Therefore PMOS can be applied in the pull-up network of pre-drive stage, as shown in Figure 5.5. Even though less strength than NMOS, PMOS helps to

reduce the response time of the pull-up network, and node x can be pulled up higher and faster with small-size PMOS. The overall performance is improved.

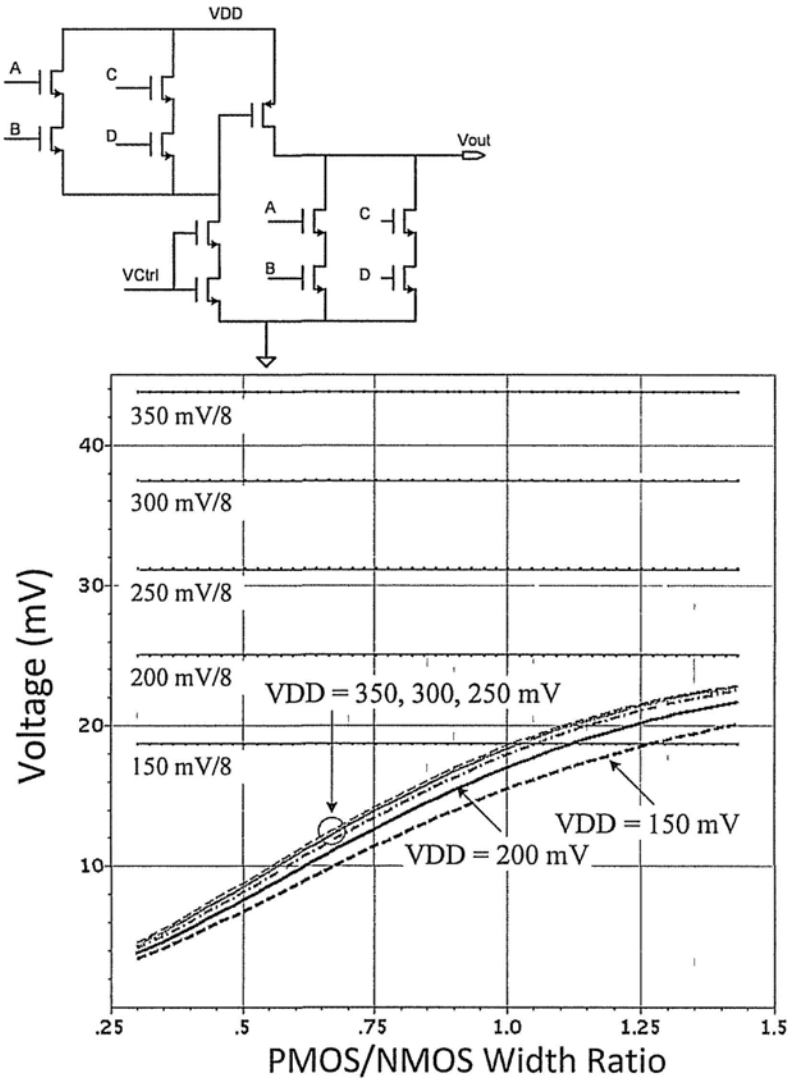


Figure 5.6 V_{OLmax} vs. effective PMOS/NMOS width ratio in different VDD of a general AOI gate of ACRL

Dynamic CMOS is another logic technique that featured with single-PMOS pull-up network. However, reliable dynamic logic must be footed by a clock-controlled NMOS transistor and extra output buffer (inverter) is necessary. In this way, the penalties are: (1) causes a longer falling delay time; (2) potential logic error due to the leakage in long evaluation period, and due to charge sharing; (3) more complex than pure combinational logic. Subchapter 6.2.1 will give comparison and discussion.

5.1.3 Design Methodology

To design an ACRL logic cell satisfies certain timing or power specification, a design flow with iteration is necessary. The selection of the width of the transistors, and the tuning of node x are the main factors to adjust power consumption. Big size input network and the active load reduce the resistance of the branch from VDD to ground, and therefore increase the leakage current and power. Tuning the node x can control the current at output stage. The level of the node x also reflect the current in pre-drive stage. Refer to timing, it depends on how fast the pull-down network can

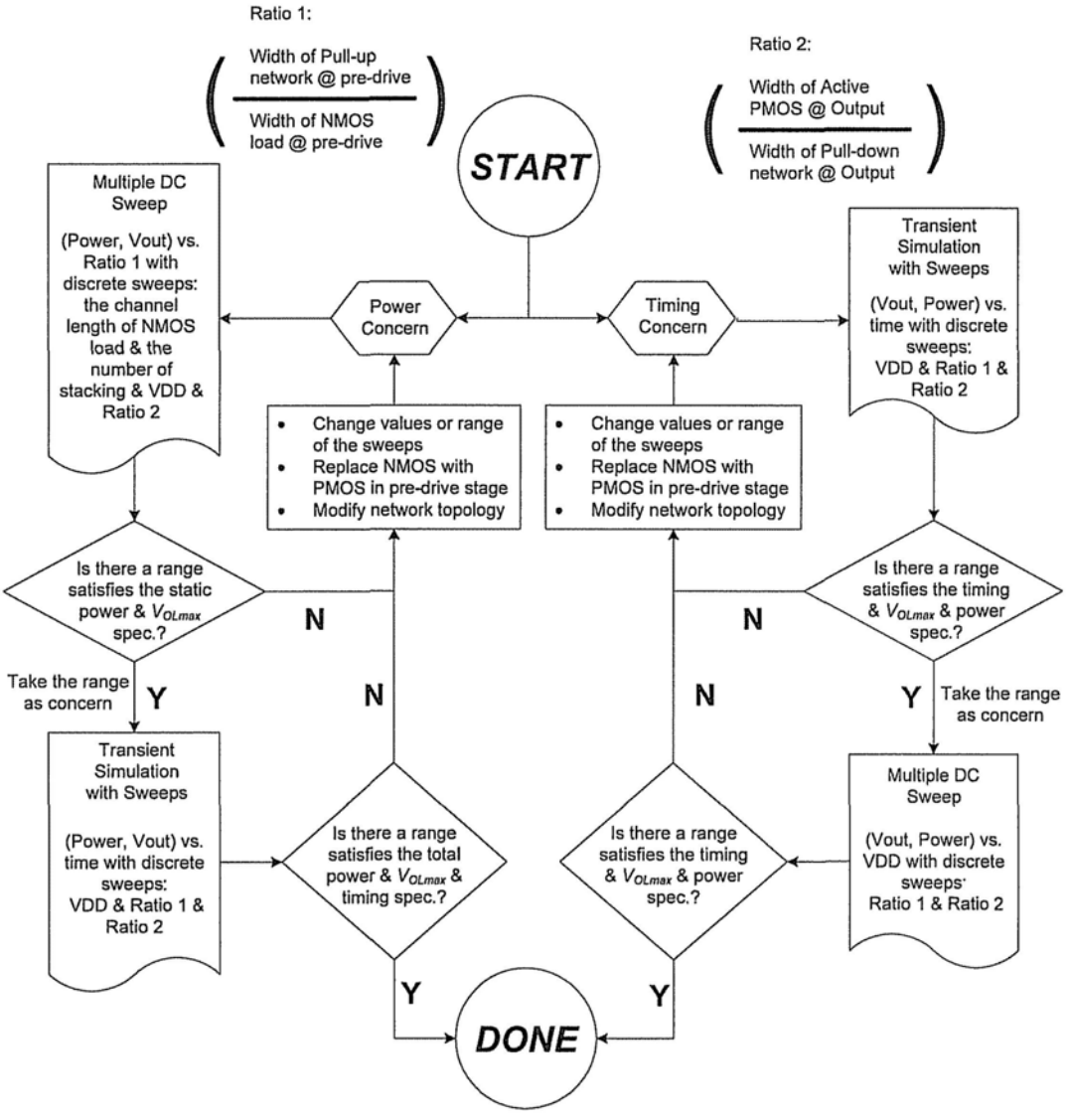


Figure 5.7 Design flow of ACRL logic cells with certain spec.

sink the output and how fast the pull-up PMOS can charge up the port. There is also a trade-off between rising and falling delay. By adjusting the PMOS to NMOS width ratio, or by changing the response speed and high level at node x, the delay time of rising and falling can be changed. Furthermore, the stacking number of transistor in the NMOS load, or the pull-up-network-to-NMOS-load strength ratio, affects the high level and response speed at node x.

As the process is fixed, the above factors are the main consideration in IC design. Based on this discussion, a design flow for ACRL cell design is proposed as shown in Figure 5.7. The design flow is a basic method to accomplish a subthreshold ACRL design, additionally the input signals transition and the output capacitive load condition should be all involved in the flow.

5.2 Logic Cell Designs

By the methods of improvement mentioned above, many logic cells can be developed in this style for a better performance in subthreshold design, especially the AOI structure, sequential cell with combinational inputs, multiplexers and magnitude comparators. Generally the topology of a logic cell is achieved by stacking transistors in the network's branch or parallelizing branches based on the inversed function. In this sub chapter several ACRL logic cells are proposed, based on the requirement of the baseband processor, the subthreshold design example, and the logic functions of general usage.

5.2.1 XOR/XNR Logic Gates

Exclusive OR and Exclusive NOR gates are the popular logic functions in most of the digital design, especially in DSP or blocks with calculation. In the design of 2-input XOR gate, we notice that the output is HIGH only when the two inputs are in

different logic value. Therefore in ACRL design, the pull-down network at output stage should be ON when the two inputs are not different, that means, the same. Figure 5.8 is the structure of ACRL 2-input XOR gate. Input A and B can be the same with logic LOW or HIGH, so there are two branches that short the output to ground with comparatively higher ON leakage current.

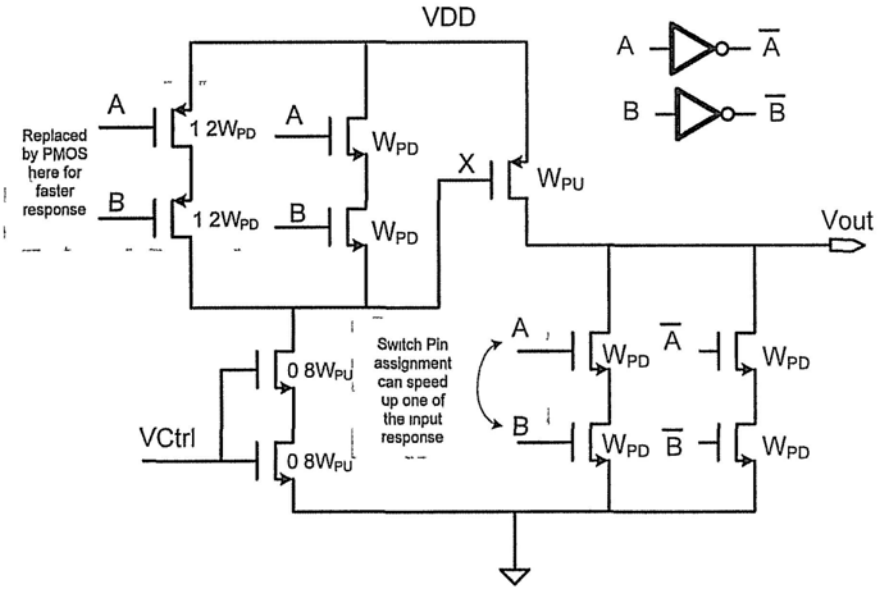


Figure 5.8 2-input XOR gate of ACRL style

The design in Figure 5.8 is a delay balanced topology. When the input is assigned to the NMOS which is close (drain connected) to Vout, it is assigned to the transistor in pull-up network that is far from the control signal, node x. Vice versa, the assignment is just flipped. If the design is aimed to have a faster response on input B, the pin assignment of A and B can be switched in the pull-down network. According to the design flow in Figure 5.7, Ratio 1 and Ratio 2 are combined into the ratio of the rising strength for Vout, to the falling strength for Vout. So the timing transient analysis is based on the ratio of W_{PU} to W_{PD} in Figure 5.8. And Figure 5.9 shows the result of the simulation, it is indicated that a bigger W_{PD} is not able to obviously speed up the falling motion at Vout. As the W_{PD} is close to W_{PU}, the 2-input XOR

design has a comparatively steady and overall shorter delay. So for the consideration of process variation, the intermediate value between 0.9 and 0.65 is selected (0.8) as the value of W_{PU} / W_{PD} in the design with W_{PU} is 320nm, and the power consumption can be lower.

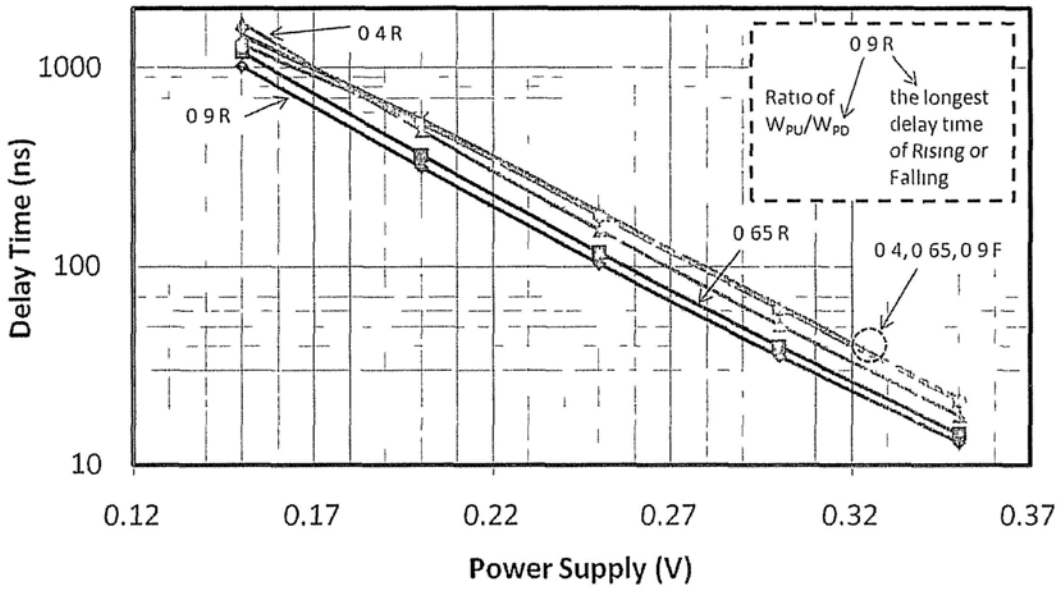


Figure 5.9 XOR timing transient simulation vs. VDD with different ratio

The 2-input XNR gate is designed as the way of building up the 2-input XOR gate. Figure 5.10 is the structure with two PMOS transistors in the pull-up network at pre-drive stage, similarly they are for faster response.

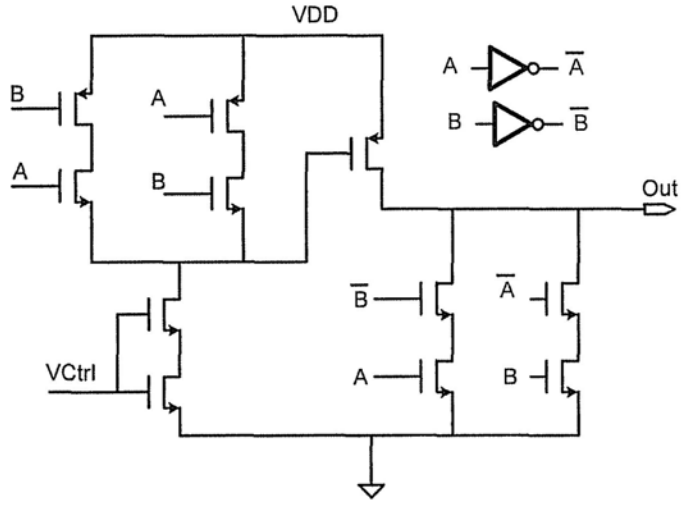


Figure 5.10 2-input XNR gate of ACRL style

The 3-input XOR gate design requires more considerations. The two transistors of NMOS load are adjusted by increasing the channel length. As the pull-down branch is now stacked by three NMOS transistors, this is a method to help node x to rise up to a higher level when the output is about to fall. To balance the delays to different inputs, the input assignment is shown in Figure 5.11. To avoid the delay of input inversion of B and C, input B and C are assigned to PMOS transistors on the top, far from node x.

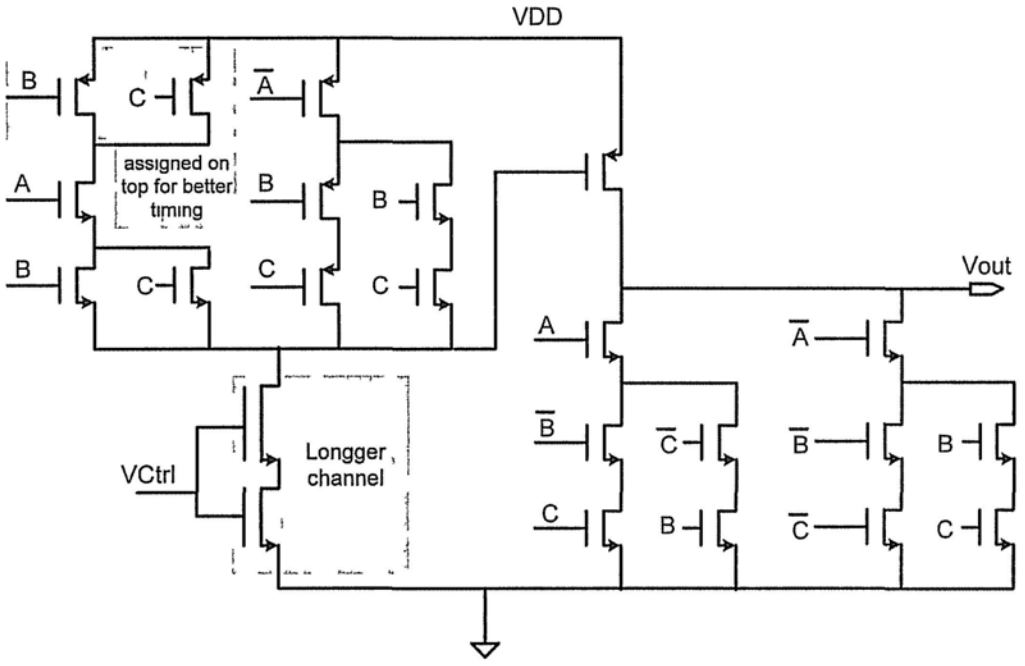


Figure 5.11 3-input XOR gate of ACRL style

5.2.2 Sequential Cell Design

Registers and D flip flops are taking an important part in digital design. For the example design baseband processor, there are hundreds of D filp flops inside. In subthreshold operation, fatal timing violation can occur when the clock-to-Q delay and setup time become very long. Figure 5.12 is the proposed ACRL D latch and DFF circuit for a lower setup time and CK-to-Q delay. Different from the

conventional DFF design, as shown in Figure 5.12, the effective delay path is reduced from the red circled TG(transmission gate)-inverter-inverter, to equivalently one 2-input-NAND gate with the short delay (in red), which is approximately just more than one inverter delay.

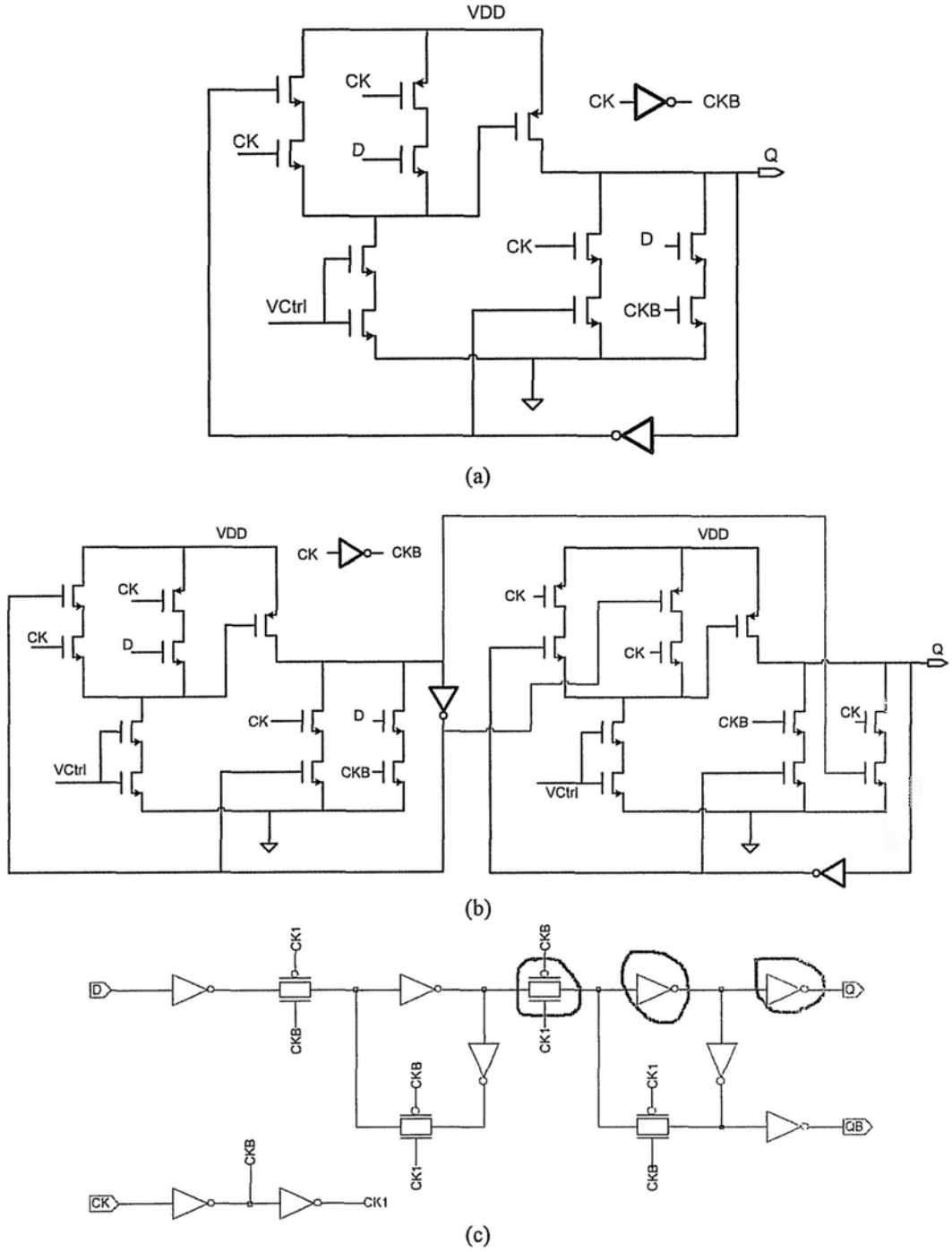


Figure 5.12 ACRL sequential cells (a) D latch; (b) D flip flop. And (c) conventional DFF

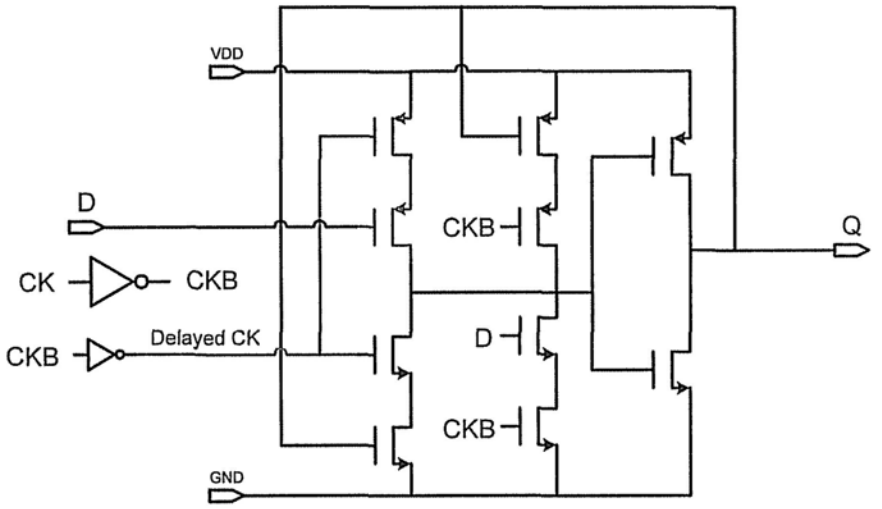


Figure 5.13 Simple and optimized D latch for shorter setup time

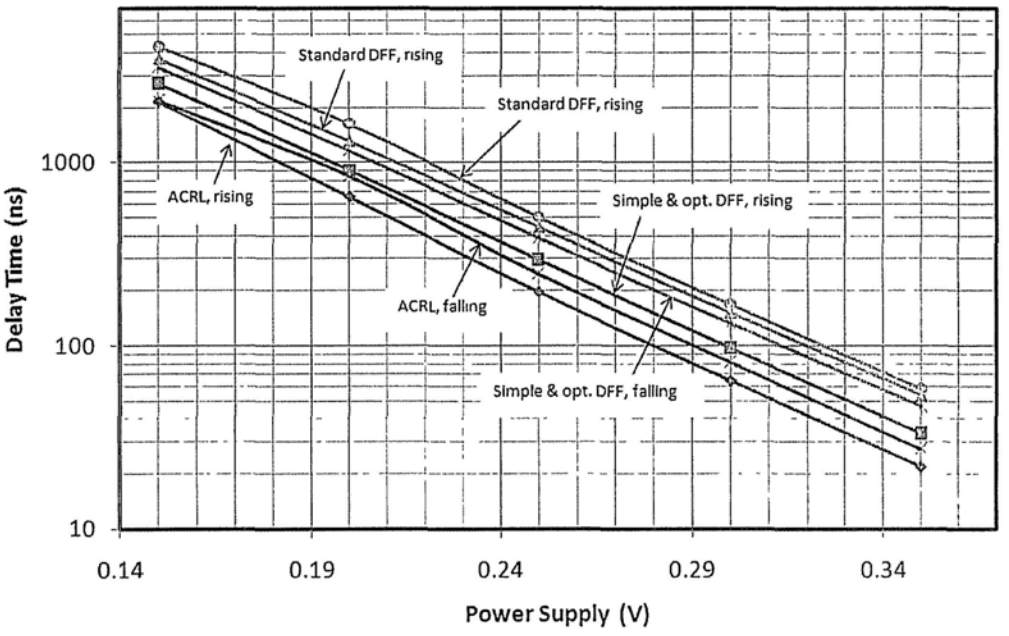


Figure 5.14 Comparison of CK-to-Q timing among three styles of DFF

To remove the static power, a simple and optimized D latch based on previous work is also proposed for subthreshold usage. Figure 5.13 is the structure of the simple D latch. Most cases of timing violation in subthreshold sequential cells are due to the setup timing error, it is hardly to violate the hold time constrant. Generally when new value is assigned by previous DFF, the logical propagation will take a long time to reach the D input of the DFF at this stage. Therefore in Fugure 5.13, delayed

clock is generated to reduce the setup time requirement of '0' D input, by keeping the branches in red active for a short time after the rising edge.

Figure 5.14 is the comparison of the delay timing of the ACRL DFF, the optimized DFF composed of the D latch in Figure 5.13, and the standard DFF cell based on the target process. In the comparison, ACRL DFF design has a much better timing in both rising and falling delay, approximately 40~60% of the delay of standard DFF cell.

5.2.3 Comparators and AOI Case Selector

In the baseband design, logic comparison is frequently required in the command verification, parameter matching and magnitude comparison in PIE decoding. The ACRL and-or-inverter structure is very suitable in designing the n-bit comparator of equality. In conception, non equality only occurs when at least in one bit position the related binary values from the two vectors are not equal. That means, given n bit vector A and B are selected to compared, then the inversed true value of the equality F can be express in Equation (5.1):

$$\begin{aligned}\overline{F} = {} & A_{n-1}\overline{B_{n-1}} + \overline{A_{n-1}}B_{n-1} + A_{n-2}\overline{B_{n-2}} + \overline{A_{n-2}}B_{n-2} \\ & + \ldots + A_1\overline{B_1} + \overline{A_1}B_1 + A_0\overline{B_0} + \overline{A_0}B_0 \end{aligned} \quad (5.1)$$

If $\overline{F}$ is '1' in true value, F is '0' in true value, otherwise, F is kept at logic HIGH. Therefore in ACRL design, the equality comparator is implemented by giving the ON branches that represent the non equality at every bit positions, to sink the output to ground. Figure 5.15 (a) and (b) are the ACRL 2-bit and 4-bit equality comparator as examples.

The equality comparator with ACRL style simplifies the logical propagation time and then a shorter delay is obtained. Especially the bit size ((effort of calculation) is

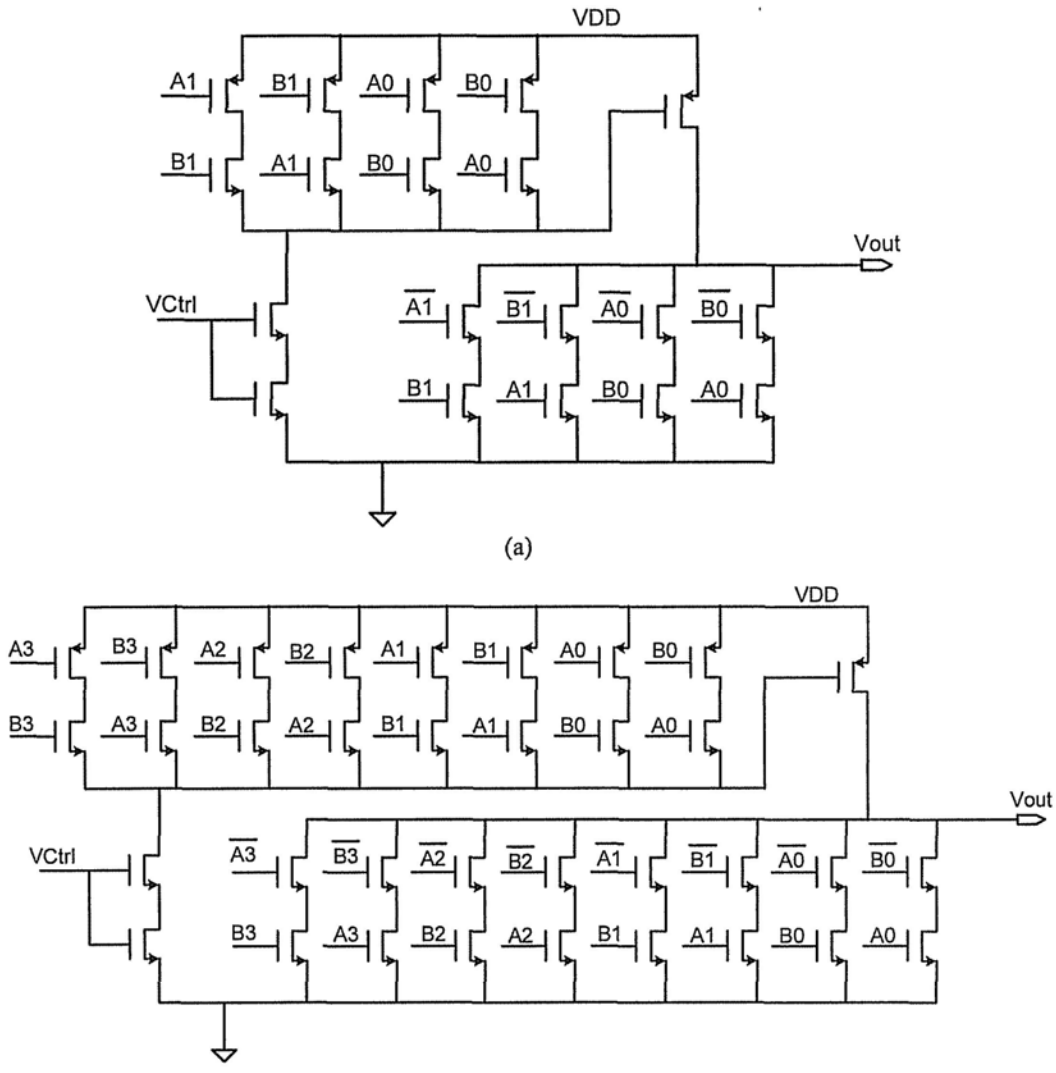


Figure 5.15 ACRL equality comparator (a) 2-bit; (b) 4-bit

bigger and more branches are required. Figure 5.16 is the comparison of ACRL 4-bit equality comparator and the static CMOS one (composed of 2-input XOR and 4-input NOR gates). The ARCL 4-bit equality comparator has a 50~70% timing improvement than the conventional standard style. It is a great improvement in subthreshold operation.

Besides equality comparison, this ACRL structure can also be utilized as multiple case selector or determiner. This selector and determiner are frequently required in this baseband processor design, especially in the state machine of the *Command*

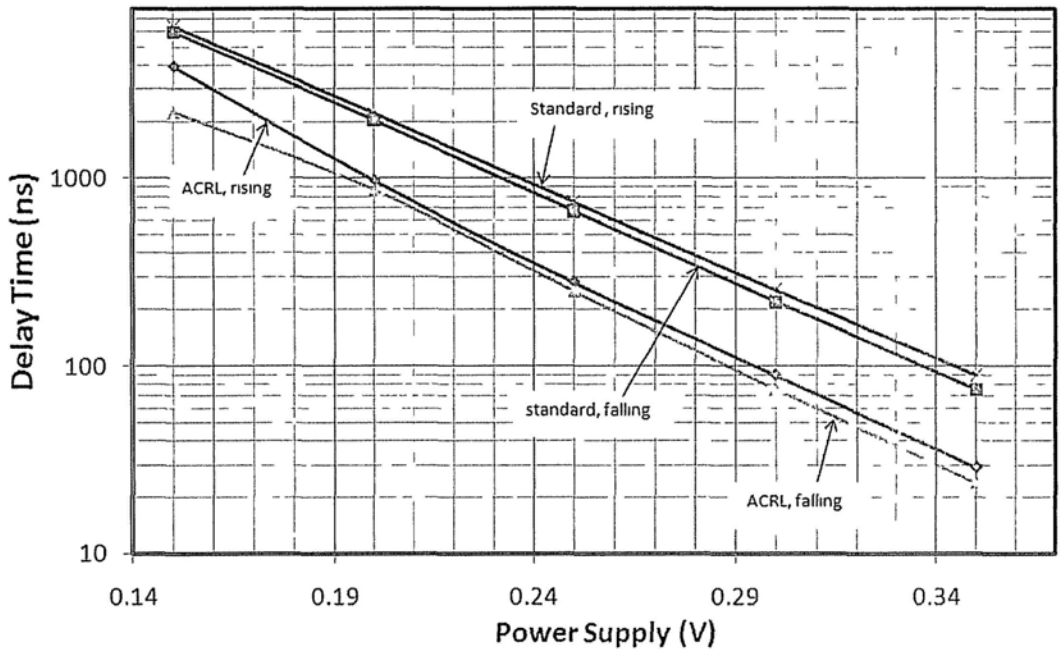


Figure 5.16 Timing comparison of 4-bit Equality Comparator between ACRL and standard structure

```
if ( (int_reg [9 8]==2'b10 & SL | int_reg
[9:8]==2'b11 & ~SL) | ((int_reg[7:6]==2'b00 &
S0_flag!=int_reg[5]) | (int_reg[7:6]==2'b01 &
S1_flag!=int_reg[5]) | (int_reg[7:6]==2'b10 &
S2_flag!=int_reg[5]) | (int_reg[7:6]==2'b11 &
S3_flag!=int_reg[5])))
Begin
    Statement
end
```

Signal Name	the True Value of	Signal Name	the True Value of
A5	int_reg [9 8]==2'b10	B5	SL
A4	int_reg [9 8]==2'b11	B4	SL
A3	int_reg[7 6]==2'b00	B3	S0_flag!=int_reg[5]
A2	int_reg[7 6]==2'b01	B2	S1_flag!=int_reg[5]
A1	int_reg[7 6]==2'b10	B1	S2_flag!=int_reg[5]
A0	int_reg[7 6]==2'b11	B0	S3_flag!=int_reg[5]

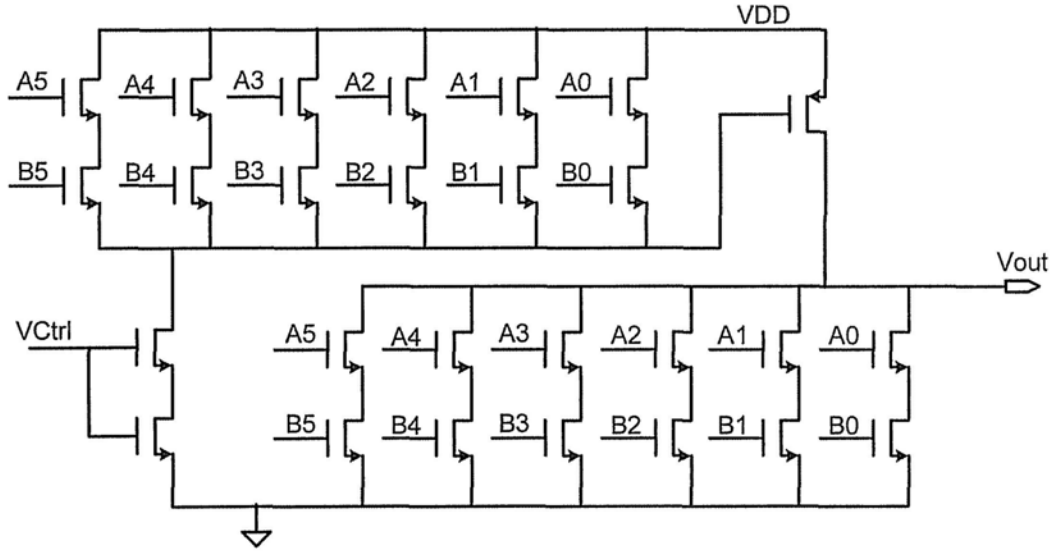


Figure 5.17 ACRL implementation of IF statement in Verilog code

Receiving and *Command Processing* block (like the AND-OR operation on the critical path in Figure 4.7). They are also important in general digital circuit designs.

Here we take an IF statement as example. It is in the Verilog coding of the state machine in *Command Processing* block, and it is on the critical timing path. As shown in Figure 5.17, the IF statement requires a parallel judgement on whether the statement should be executed. In this condition, the 6-branches ACRL AOI logic cell is perfectly matched to the Verilog code. According to different case selection or determination, ACRL circuits with relative branches, sub-branches and stacked transistors, can be developed to speed up the critical timing paths.

5.2.4 Magnitude Comparators

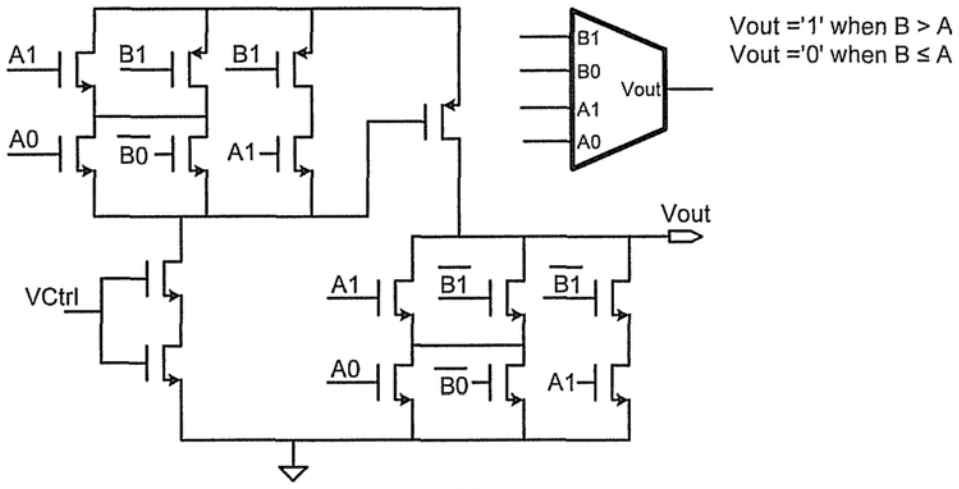
In the PIE Decoder of the baseband processor, magnitude comparators are used to identify the data symbols and the header of the command when the tag starts to receive modulated RF signal. Magnitude comparators are also applied in digital circuits like DSP modules and communication block. ACRL structure is very suitable for fast GT (greater than) comparators, and with proper configuration, fast (relatively in subthreshod) n-bit magnitude comparators and fast binary adder/subtractor can be built up with ACRL GT comparators.

To define the topology of ACRL GT comparator, it needs to find out in which input pattern, the output is LOW. If a GT comparator is used to determine whether B is greater than A, then in ACRL design we should add branches which stand for "B is less or equal (LE) to A" in the pull-down network. Like the 2-bit GT comparator in Figure 5.18(a), if A is "11", no matter how B is assigned, the output should be grounded. This also happens when A1 is '1' and B0 is '0', or A0 is '1' and B1 is '0', or B is the minimum value "00". Additionally, output is grounded when A is "10"

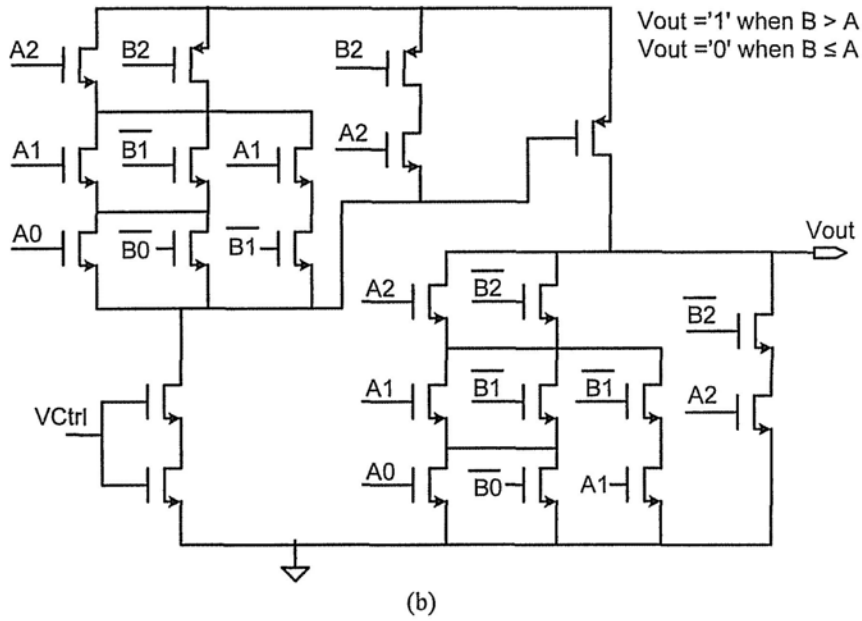
and B is "01". In logic equation, the inversed value of output F is pursued and simplified like Equation (5.2):

$$\begin{aligned}\overline{F} &= A_1A_0 + A_1\overline{A_0}(B_1\overline{B_0} + \overline{B_1}B_0 + \overline{B_1}\,\overline{B_0}) + \overline{A_1}A_0(\overline{B_1}B_0 \\ &\quad + \overline{B_1}\,\overline{B_0}) + \overline{B_1}\,\overline{B_0} \\ &= (A_1 + \overline{B_1})(A_0 + \overline{B_0}) + A_1\overline{B_1} \end{aligned} \tag{5.2}$$

Using this method of branch arrangement, Figure 5.18(a) and (b) give the proposed ACRL 2-bit and 3-bit comparator respectively.



(a)



(b)

Figure 5.18 ACRL GT comparator (a) 2-bit; (b) 3-bit

4-bit magnitude comparator is popular with in industry and it is the basic unit to form longer-bit-size magnitude comparators. Figure 5.19(a) is the standard 4-bit magnitude comparator, and Figure 5.19(b) is the proposed ACRL 4-bit GT comparator. It utilizes ACRL 2-bit GT comparator as the operation unit, and output '1' when *DataI* is greater than *DataII*.

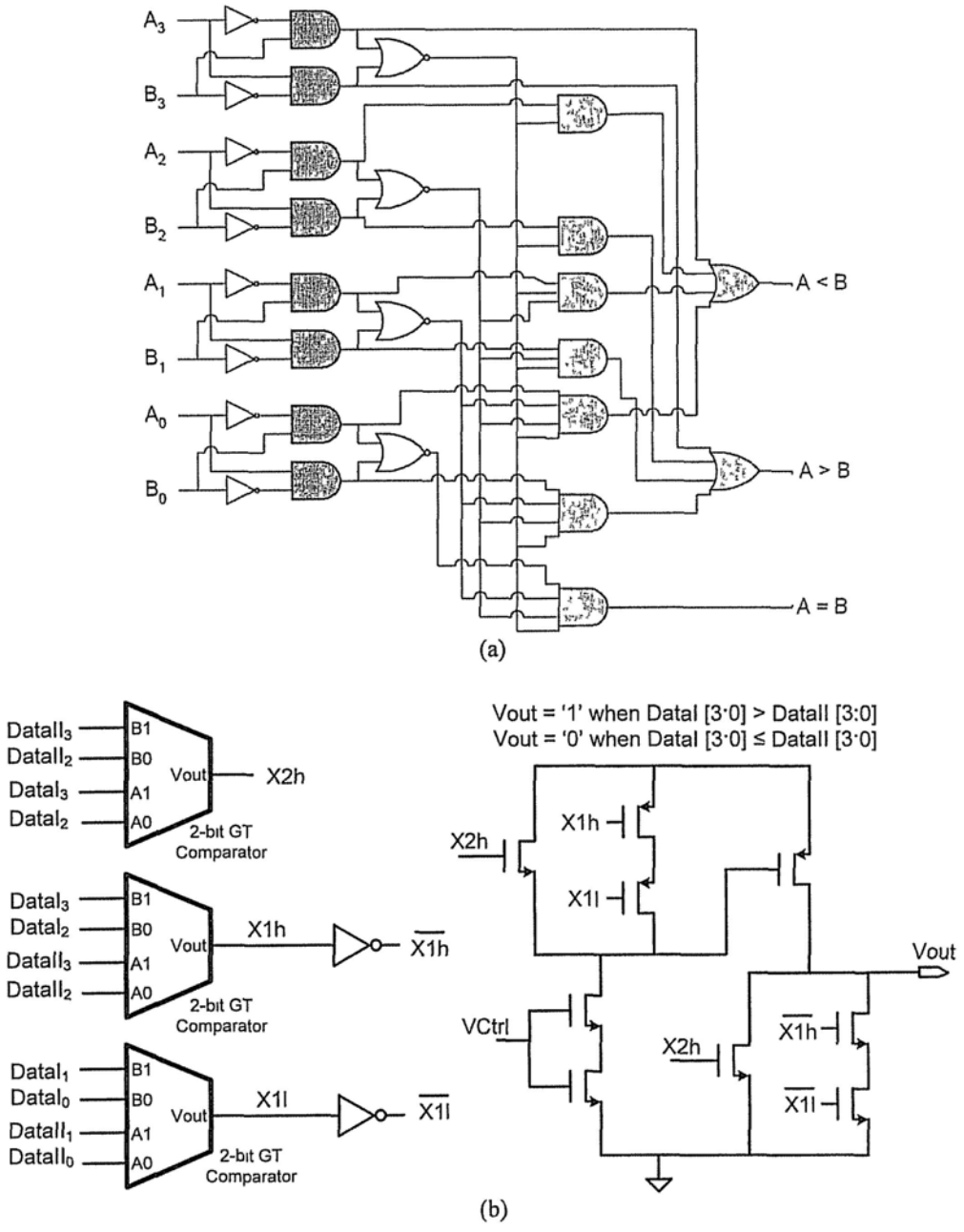


Figure 5.19 ACRL 4-bit GT comparator (a) standard static CMOS; (b) ACRL

The comparison results in a good reduction of delay time. As shown in Figure 5.20, the worst-input-case rising and falling delay of ACRL 4-bit GT comparator is 40~70%.

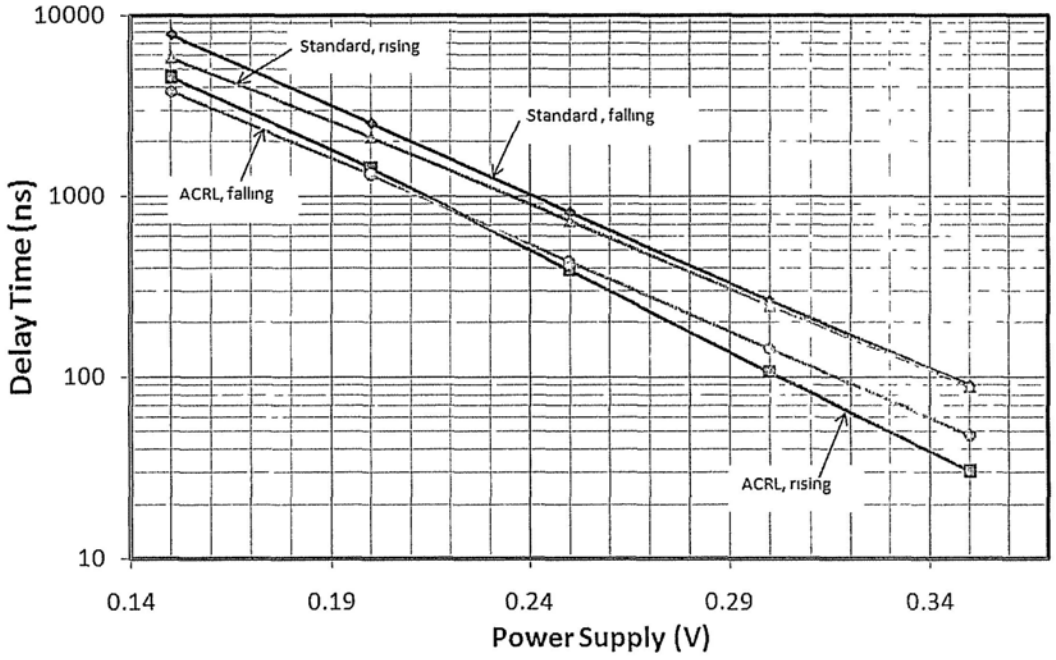


Figure 5.20 Comparison of standard and ACRL 4-bit GT comparator

5.3 Compound Logic & Design for Test

To verify the functionality and performance of the proposed ACRL circuits, some of these custom cells are implemented in ASIC chip for measurement. The cells are applied in chains or rings for a better testability. And based on the actual application of the example BBP design in this thesis, a custom simplified PIE decoder is also implemented on silicon. It is built up with mainly ACRL cells and executes the PIE data symbol decoding like the PIE Decoder design in Chapter 4, except the function of header detection and the interactive module with the Command Receiving block.

5.3.1 ACRL Cells in Chain & Ring

The measurement of the logic cells requires chain or ring circuity to have a good readability on the oscilloscope. Among the layouts of the proposed custom ACRL cells in sub chapter 5.2, the 2-input XOR gate, the 2-bit equality comparator, the 4-bit GT comparator and the D filp flop are selected to be fabricated in the target process, as well as the related standard cells for comparison in measurement. The

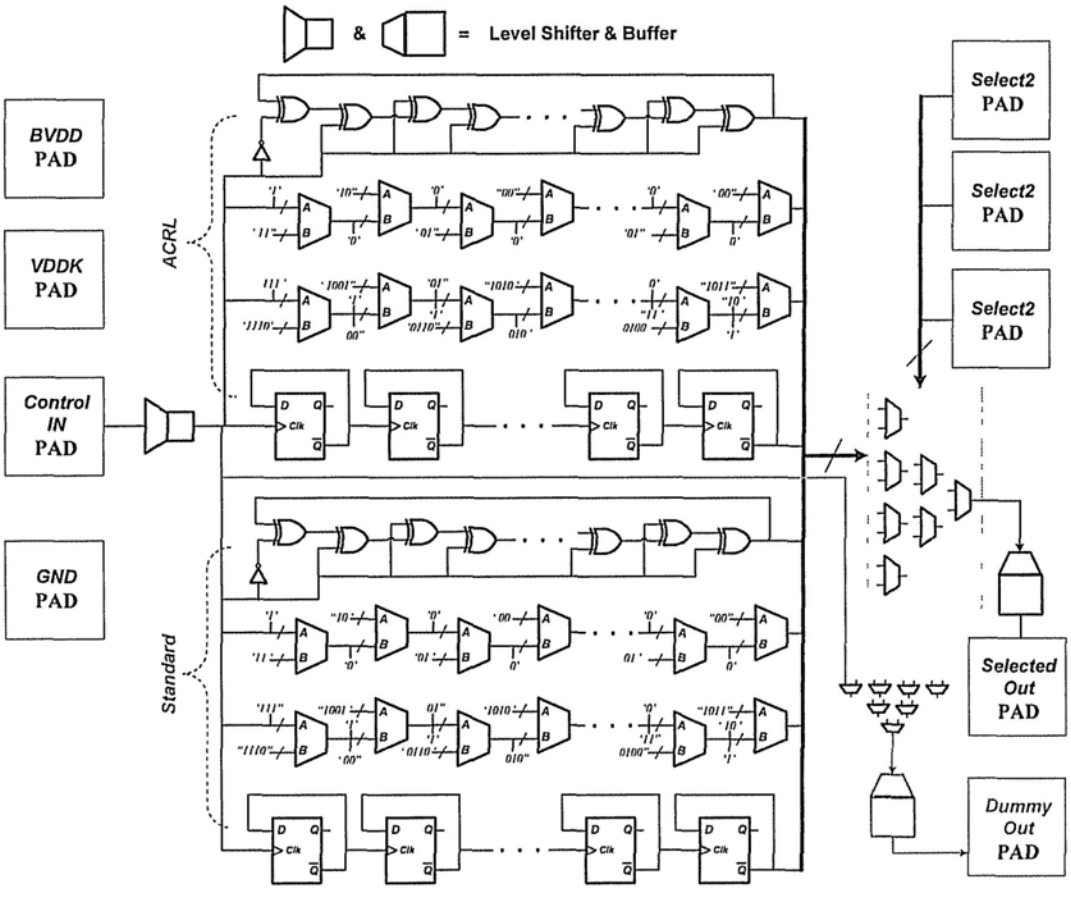


Figure 5.21 Chain & ring structure for testability

corresponding standard logic gates are also implemented for comparison. Figure 5.21 is the circuit that showing the ring and chain connections of these cells. The ACRL and standard 2-input XOR gates are connected in ring, 20 stage respectively. The connection is assigned as negative feedback, so the average delay t_d can be derived from the oscillating frequency when the rings are powered up, like:

$$t_d = \frac{1}{2 \times 20 \times freq} \tag{5.3}$$

Other cells are connected in serial chains, they are 16 or 8 staged that the propagation time from the input of the first stage to the output of the last stage, can be viewed and measured by the oscilloscope. A dummy input-to-output connection is utilized to reflect the delay time when the input passes level shifters, multiplexers, pads and other outer circuits. In this way the actual delay of the core circuit is measurable.

5.3.2 Complex ACRL Logic Circuits

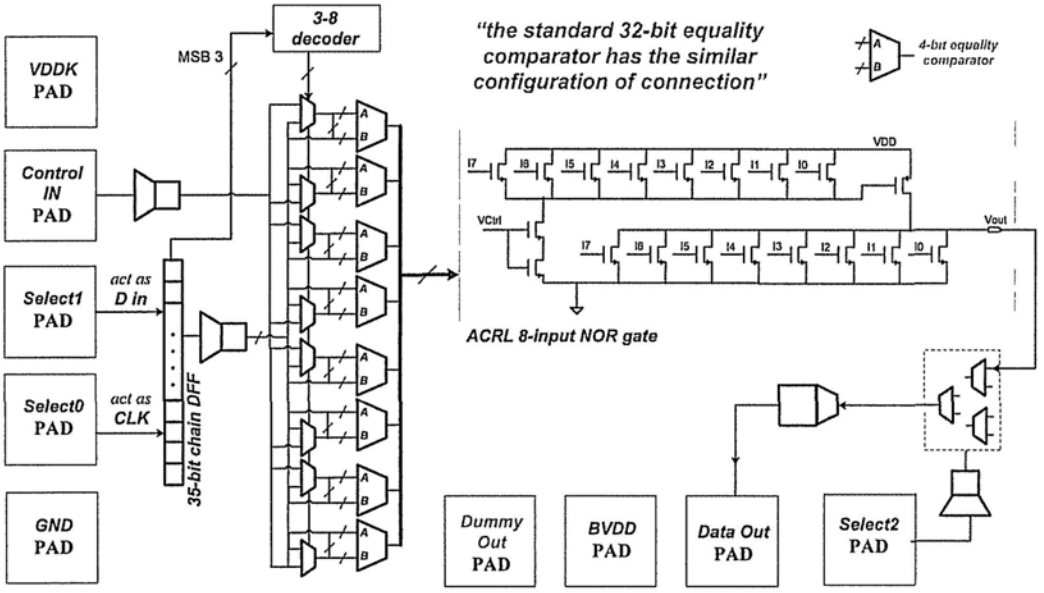


Figure 5.22 ACRL 32-bit equality comparator and test circuits

To verify the performance of ACRL cells in general usage, an ACRL 32-bit equality comparator is fabricated, as well as the static CMOS 32-bit equality comparator for comparison. The ACRL 32-bit equality comparator is composed of 4-bit equality comparator and the ACRL 8-input NOR gate, as shown in Figure 5.22. There are eight bit position in the two 32-bit inputs are selected to test the delay of the comparators, and chain DFFs are applied to input the test pattern. As shown in the same figure, the chain DFF block has 35 one-bit DFF which are serially

connected. The extra 3 bits are assigned to the 3-8 decoder to choose at which input and which bit position will be controlled by the *Control IN* PAD in measurement.

An ACRL PIE decoder prototype is also implemented in the target process to verify the performance of ACRL style in a practical digital design. The ACRL PIE decoder has the core function of PIE decoding, and covers the critical path of the full-function PIE Decoder in Figure 4.10 of Chapter 4. Only the header detection and the state interaction with the *Command Receiving* block are waved. As shown in Figure 5.23, the two DET DFFs in the Mixed Counter are composed of the ACRL D latch. As the timing condition is not tight in #2 position, the ripple sub-counter portion uses 9 SET DFFs, which are composed of the optimized DFF in Figure 5.13.

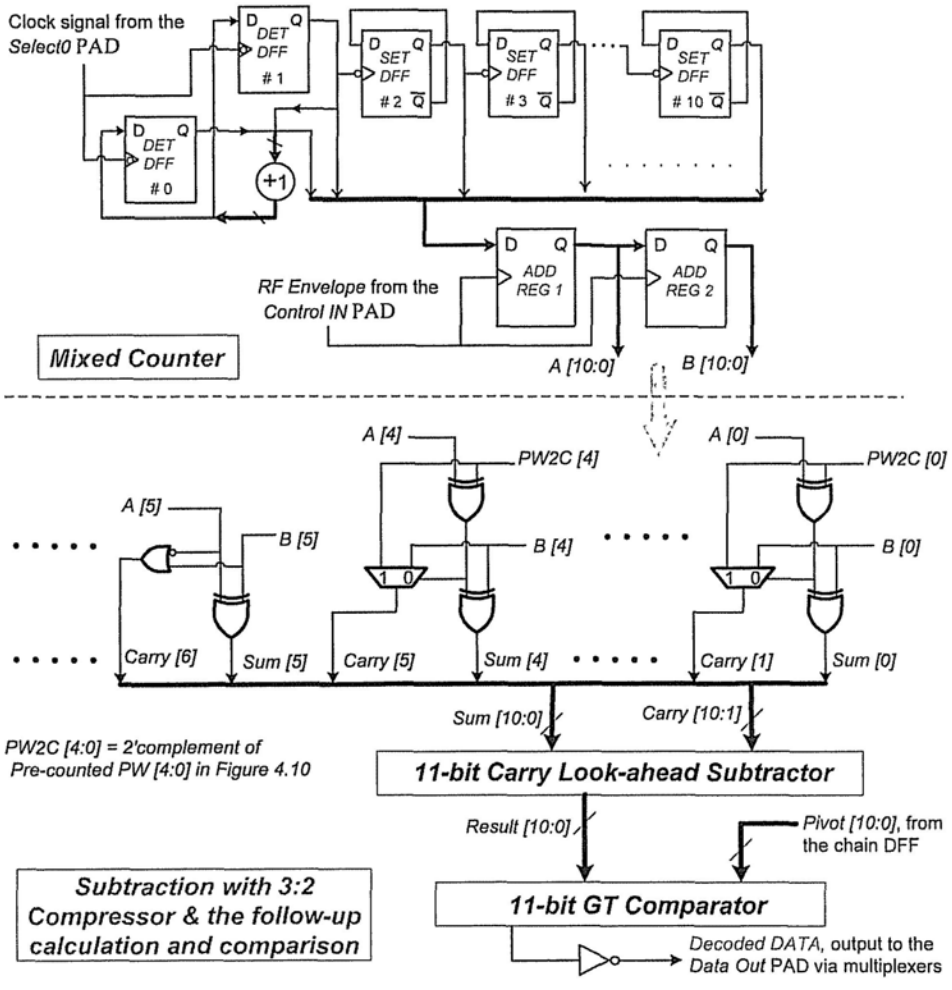


Figure 5.23 Structure of ACRL PIE Decoder

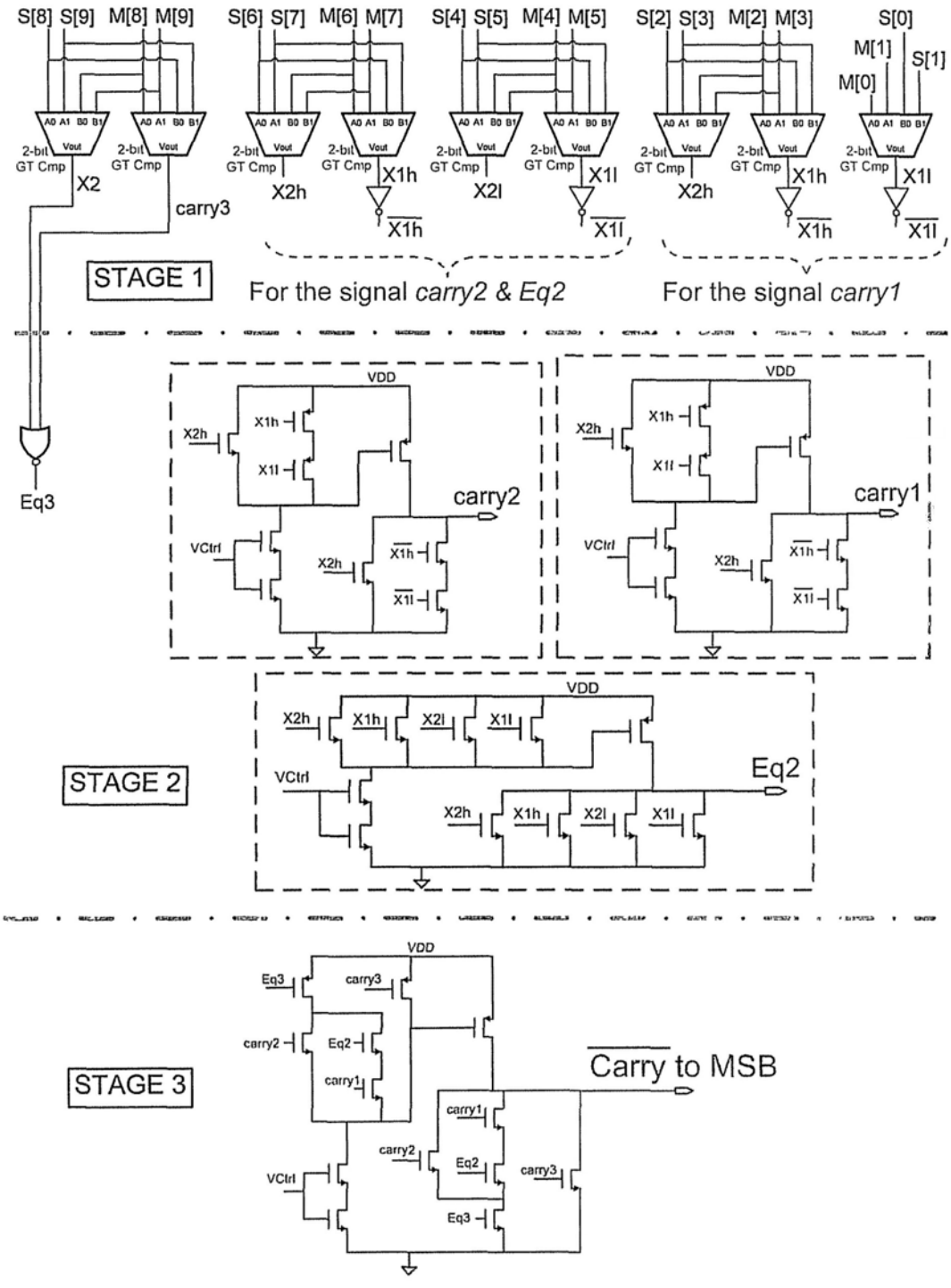


Figure 5.24 ACRL structure of carry propagation to the MSB, in the calculation of *M[10:0] - S[10:0]*

The subtract-and-add calculation in Figure 4.10 is optimized into purely subtraction in Figure 5.23. The rightmost 5 bit is 3-colum subtraction using 3:2 compressor, and the leftmost 6 bit is 2-column subtraction using 2:2 compressor. The *Sum* and *Carry*

after compression are input into the carry look-ahead subtractor as the minuend and subtrahend respectively. The carry propagation circuit for the MSB in the subtractor is a novel structure for fast speed in subthreshold operation. Figure 5.24 is the detail of the structure. The circuit is mainly formed by ACRL 2-bit GT comparators. In mathematics, carry is true when the subtrahend is greater than the minuend. Therefore in the first stage of carry propagation in Figure 5.24, 2-bit GT comparators are applied. In the second stage, for the rightmost 8 bit, the GT and equality results are come out in 4-bit size. And finally at the third stage, the inversed value of the carry for the MSB is obtained. The 11-bit GT comparator in Figure 5.23 has almost the same structure like what is shown in Figure 5.24, except that the 3-bit GT comparator and 3-bit equality comparator are applied in the leftmost 3 bit (GT comparison requires the carry from the MSB not for the MSB). They are shown in Figure 5.25. And Figure 5.26 is the layout of the PIE decoder.

The ACRL PIE decoder, 32-bit equality comparator and the cells in chain & ring are implemented together with shared pads. Chapter 6 will give the results of measurement, as well as the BBP design in Chapter 4 using only standards cells.

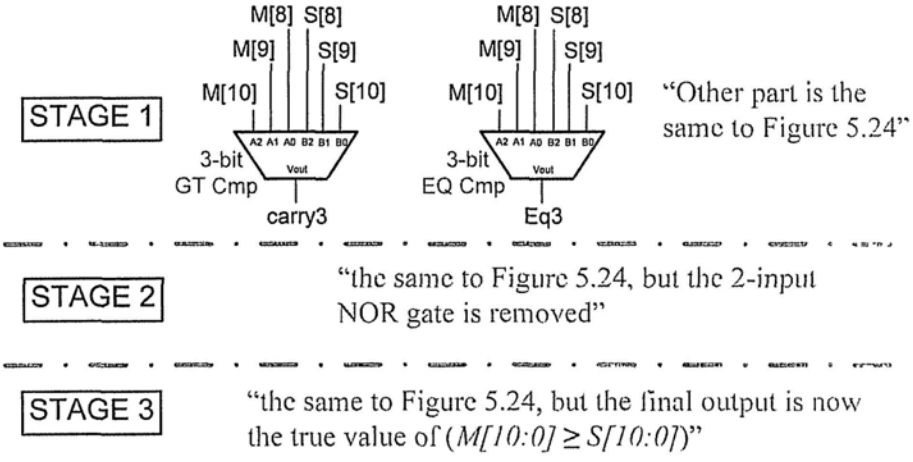


Figure 5.25 ACRL 11-bit GT comparator, with inversed-value output

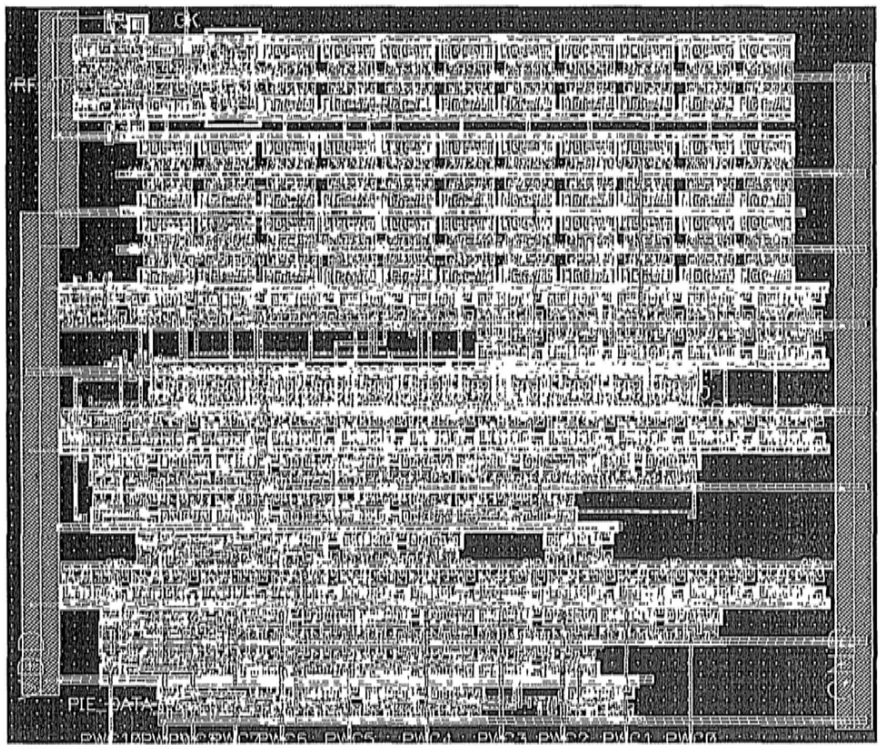


Figure 5.26 ACRL PIE decoder

CHAPTER 6 EXPERIMENTAL VERIFICATION

The ultra purpose of this subthreshold research is to verify the feasibility of the proposed subthreshold structures and topologies in system-level, gate-level and transistor-level. The verification is based on the actual design examples, from the small system: EPC C1G2 baseband processor, the complex modules: the ARCL 32-bit equality comparator and ACRL PIE decoder, to the basic logic units: the ACRL cells in chains and ring. In this chapter, the physical verifications are carried out to investigate the performance of these circuits by different methods.

6.1 Baseband Processor Verification

The baseband processor design based on EPC C1G2 protocol has been implemented in the UMC 130nm CMOS technology and UMC 90nm CMOS technology repectively. This proposed design covers most of the EPC C1G2 specified commands except the optional *Access*, *BlockWrite*, *BlockErase*, *Lock* commands and the mandatory *Kill* command. Here it is called the complex version. To verify the cooperation with other analog blocks in the passive RFID tag, the basic baseband processor that covers the commands for inventory communication (*Query*, *QueryAdjust*, *QueryRep*, *Ack* and *NAK*) has been integrated in the 180nm, 130nm and 90nm chip fabrication of the entire tag previously. Here it is called basic version.

6.1.1 Simulation and FPGA test

The verifications on the baseband processor are carried out in the two periods of pre-fabrication and post-fabrication. Before the ASIC fabrication, the baseband design has gone through digital front end to back end flow. Besides the structure optimization for subthreshold usage, as describe in Chapter 4, there is also a rule in

standard cells selection. Scripts are added to forbid the adoption of high fan-in standard cells with over-three stacked transistors. In the frequency constrain, because only the timing information at nonimal power supply is alvailable in standard cells library files. So in the design flow of logic synthesis and place & routing, the frequency mapping method besed on Equation (2.13) in Chapter 2 is adopted. If the post-layout simulation satisfy the clock frequencies after the mapping, that means the design is probably function in subthreshold region. Equation (2.13) is a rough approximation of timing, hence the setup and hold time margin are set to be quite long in the backend design, about dozens of nanoseconds.

Figure 6.1 is the waveform of post layout simulation, the simulation has covered all the data rates specify in the EPC C1G2 ptotocol. Except the exact data rates, the data rates after being mapped by Equation (2.13) are also applied.

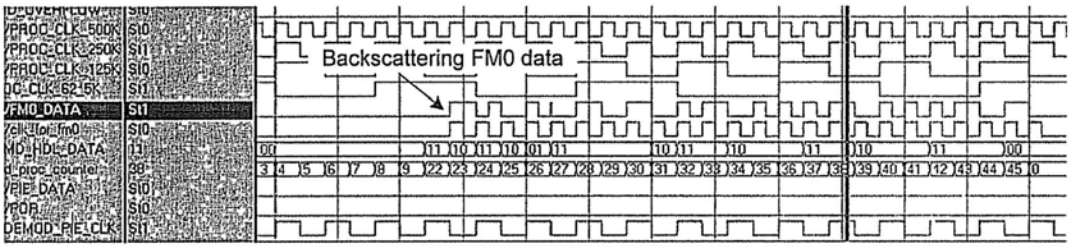


Figure 6.1 Post layout simulation waveform of BBP design

Besides the simulation by EDA tools, the BBP design is implemented into the FPGA board for logic function verification in another aspect. Figure 6.2 is the FPGA boards for logic function verification, one is implemented by the BBP design's code, and the other one is implemented by the reader's code. The designed reader can decode and verify the validability of the signal backscattered by BBP. It also send PIE encoded commands to the BBP accordingly and conduct the communication round.

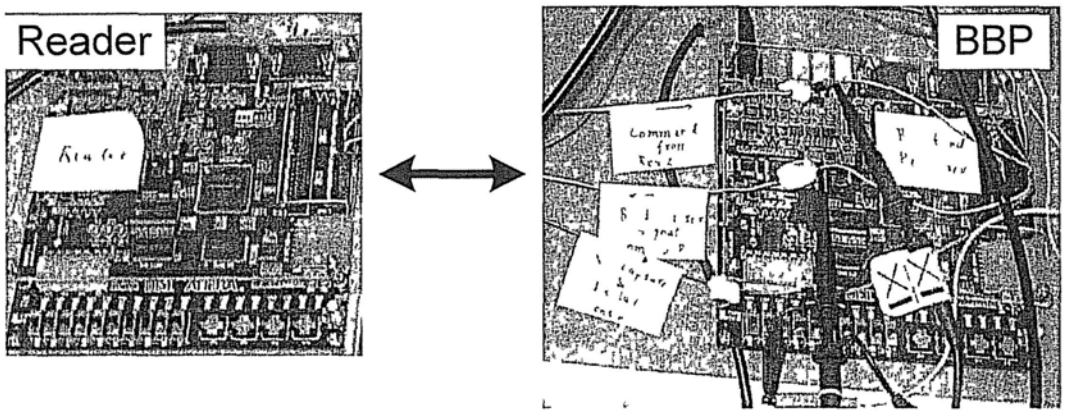


Figure 6.2 FPGA board verification of reader-and-tag (BBP) communication

Both the post layout simulation and FPGA verification have confirmed the functionality of the BBP design, enhanced the reliability of ASIC implementation

6.1.2 Chip Measurement: 130 nm CMOS Technology

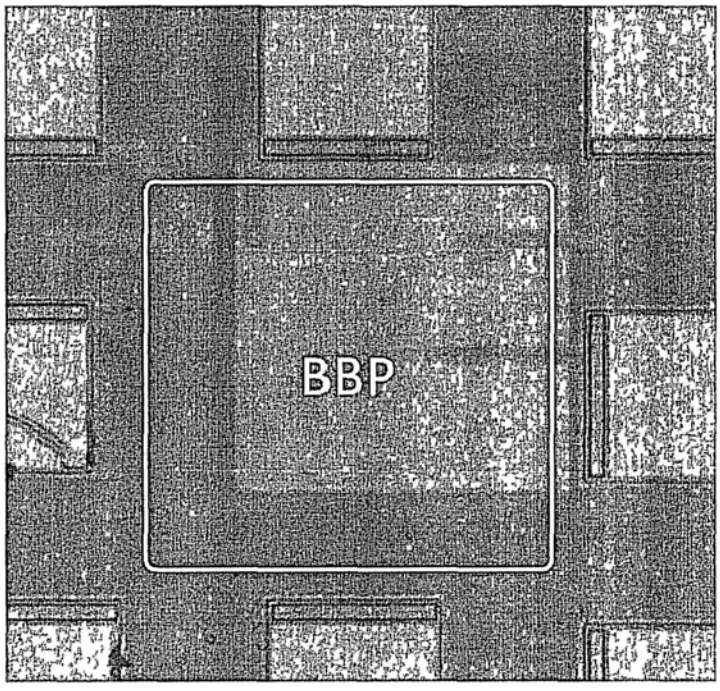


Figure 6.3 Die photograph of the BBP fabricated in 130 nm CMOS technology

Figure 6.3 is the die photo of the BBP design fabricated in 130 nm CMOS technology. In measurement the FPGA board is utilized as the reader, sending commands to the BBP on chip accordingly. The BBP indentifies and responses to the

sent command, and then the oscilloscope and the FPGA reader capture the responses for verification, as shown in Figure 6.4. A tunable square wave source (the signal generator) is also applied to act as the clock input from ring oscillator

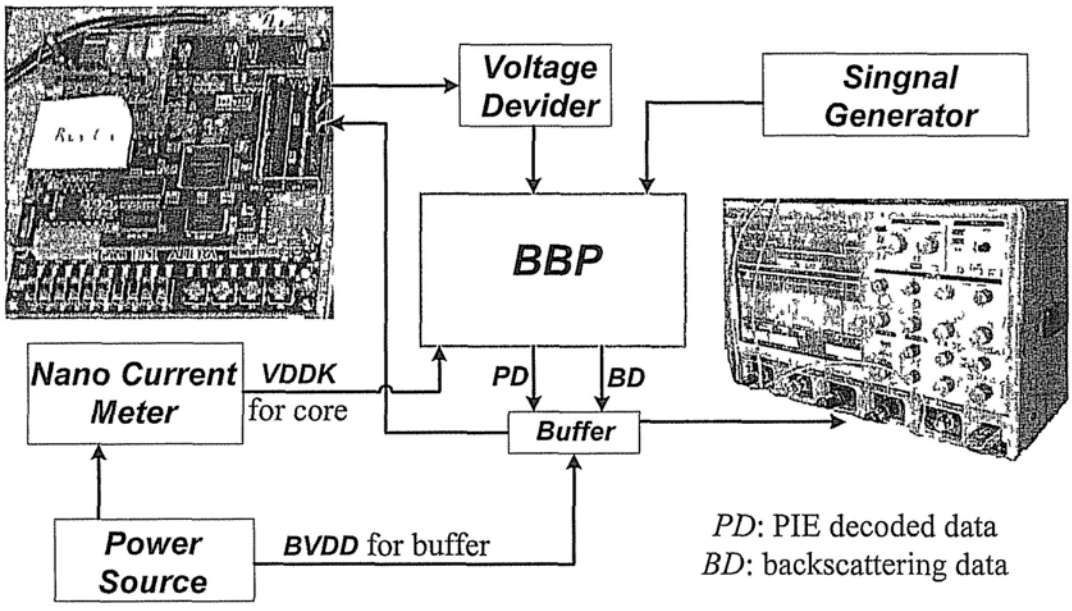


Figure 6.4 Configuration for chip measurement

In the testing, EPC C1G2 based commands were given to the chip in iteration. To inspect the tolerance and performance of the chip, for different command iterations with different *Tari*, *RTCal* and *TRCal* length, the power supply was raised from 0.25 to 1.2 V in selected points, then was reduced down to 0.25 V again. In each step, the square-wave clock input would be changed incrementally between 0.64 and 2 MHz. Meanwhile the multimeter, oscilloscope and FPGA monitor the outputs So the results of power consumption and functionality in these conditions were therefore obtained.

For all the data rate (determined by *Tari*, *RTCal* and *TRCal*) conditions, the baseband processor is able to function under the power supply from 0 33V to 1.05V. Due to different input clock frequencies and supply voltages, the fully-functioned baseband processor's power dissipation can be over 2 μW or as low as 0 238 μW.

Figure 6.5 is the captured waveform of the functionality testing of the BBP design from the oscilloscope. And Figure 6.6 is the power consumption of the processor versus supply voltage with different input clock frequencies, with the *Query*

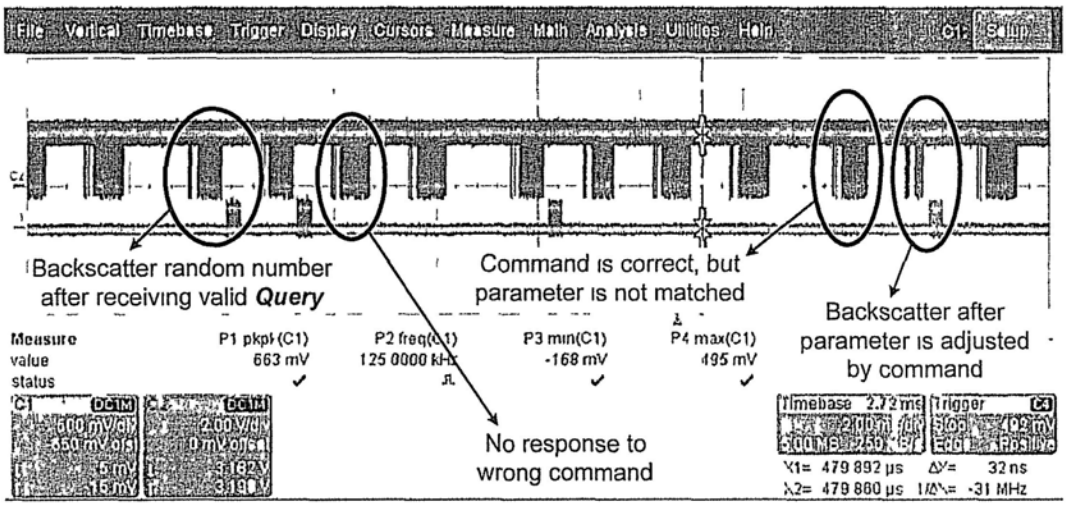


Figure 6.5 Captured waveform from oscilloscope, in the functionality test.

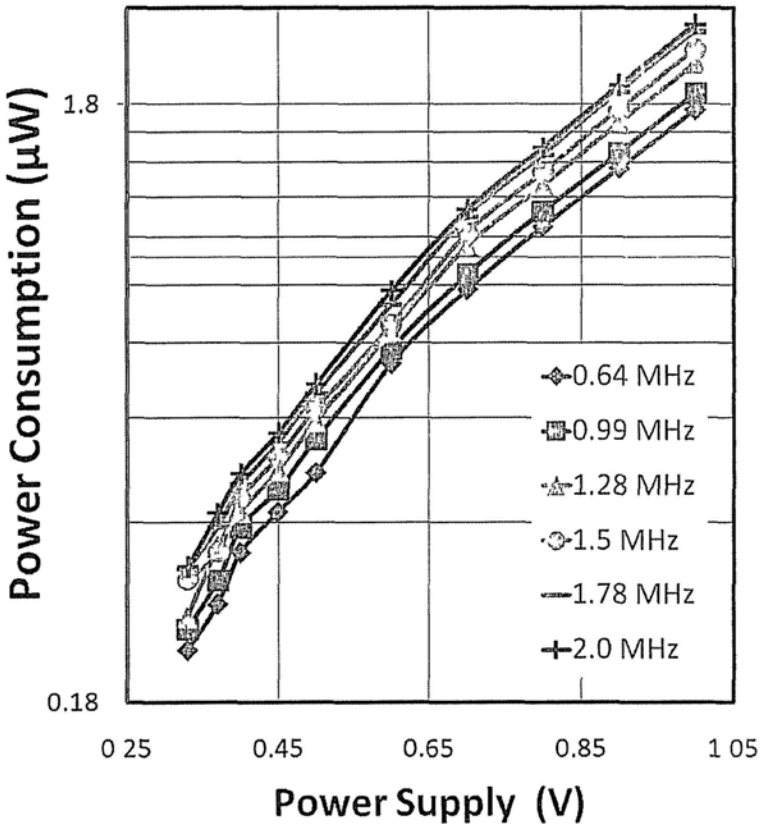


Figure 6.6 Power consumption of the BBP versus power supply with different input clock frequency

command input iteration, and followed by backscattering output from the processor (as shown in Figure 6.7).

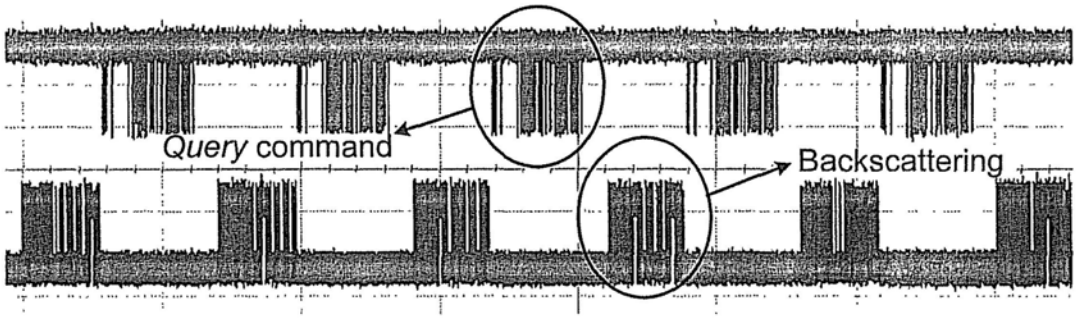


Figure 6.7 Command input iteration for power measurement

In the BBP design, the decoded data is also selected to be the output of the chip. In this way the proposed subthreshold PIE decoder in sub chapter 4.2.1 can be verified in the measurement. Figure 6.8 is the captured waveform of the decoded data as the PIE encoded command is sending to the chip.

With the structure improvement in sub chapter 4.2.1, the PIE decoder on chip can normally function with all the required data rates at a minimum supply voltage of 0.3 V. Refer to the PIE decoder design with general structure (which was fabricated in another MPW run previously), it only function as normaly at the VDD above 0.4 V. It is a great progress for the proposed design.

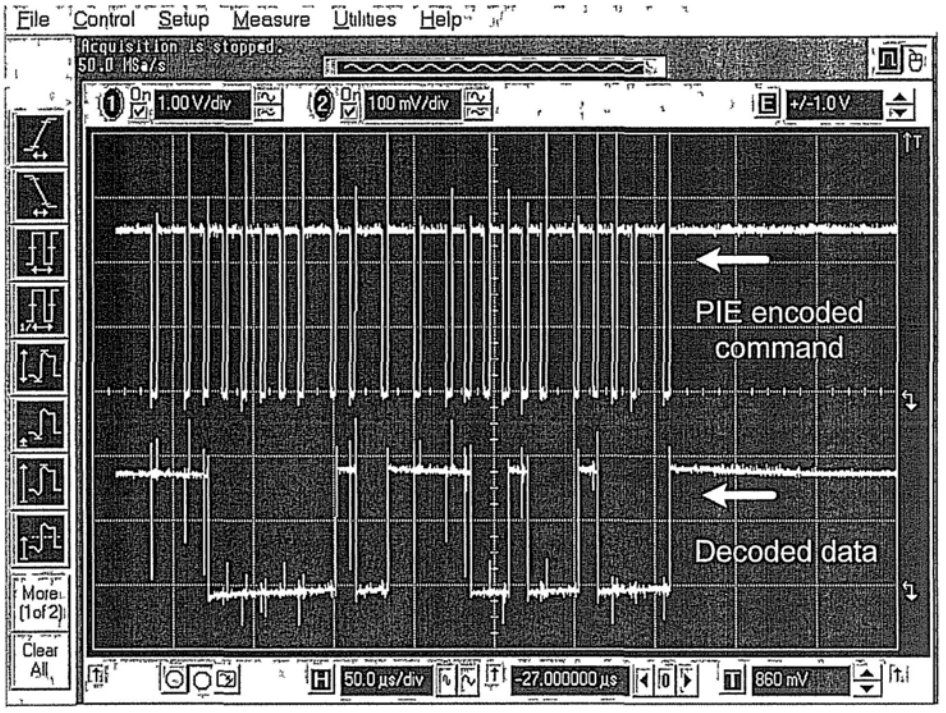


Figure 6.8 Captured waveform of PIE decoding verification

Figure 6.9 is the figure of decoding margin versus *Tari*, similar to Figure 4.14. The maximum and minimum numerically-possible values of the decoding margin for data-0 and data-1 were plotted in solid lines. The smallest decoding margin for every selected value of *Tari* in post-layout simulation, when VDD was varied from 0.3 V to 0.7 V, and all the PVT corners were run, was also picked out and plotted. And finally the decoding margins from measurement results are added to the figure. In this way the reliability of the proposed PIE decoder in subthreshold operation can be highly ensured, as the values of decoding margin are all above zero.

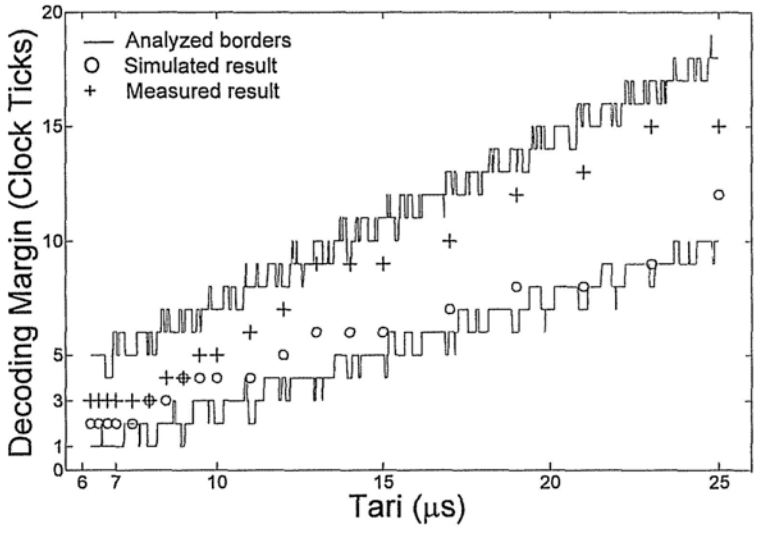


Figure 6.9 The analyzed, simulated and measured Decoding Margin vs. *Tari*

6.1.3 Chip Measurement: 90 nm CMOS Technology

In the purpose to investigate the feasibility of the proposed BBP design in deep submicron technologies. The BBP design is also fabricated in the 90 nm and the 180 nm CMOS technology. Figure 6.10 is the die photo of the BBP design of 90 nm CMOS technology. The design core is covered by dummy metal.

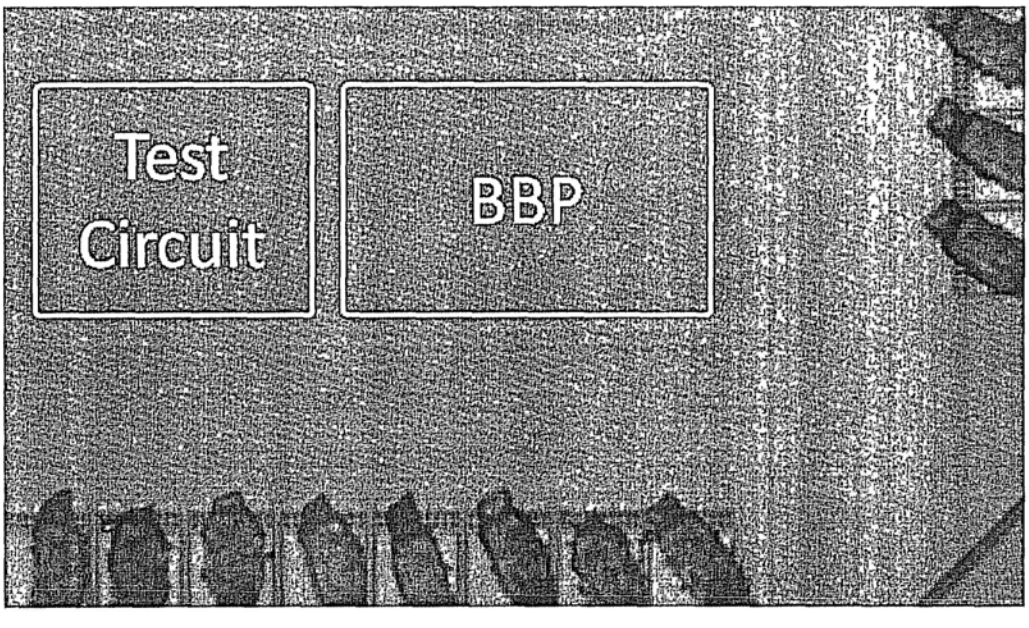


Figure 6.10 Die photograph of the BBP fabricated in 90 nm CMOS technology

For all the data rate (determined by *Tari*, *RTCal* and *TRCal*) conditions, the baseband processor is able to function under the power supply from 0.33V to 0.7 V. Due to different input clock frequencies and supply voltages, the fully-functioned baseband processor's power dissipation can be near 2 μW or as low as 0.179 μW. Similar to Figure 6.6, Figure 6.11 is the power consumption of the processor versus supply voltage with different input clock frequencies.

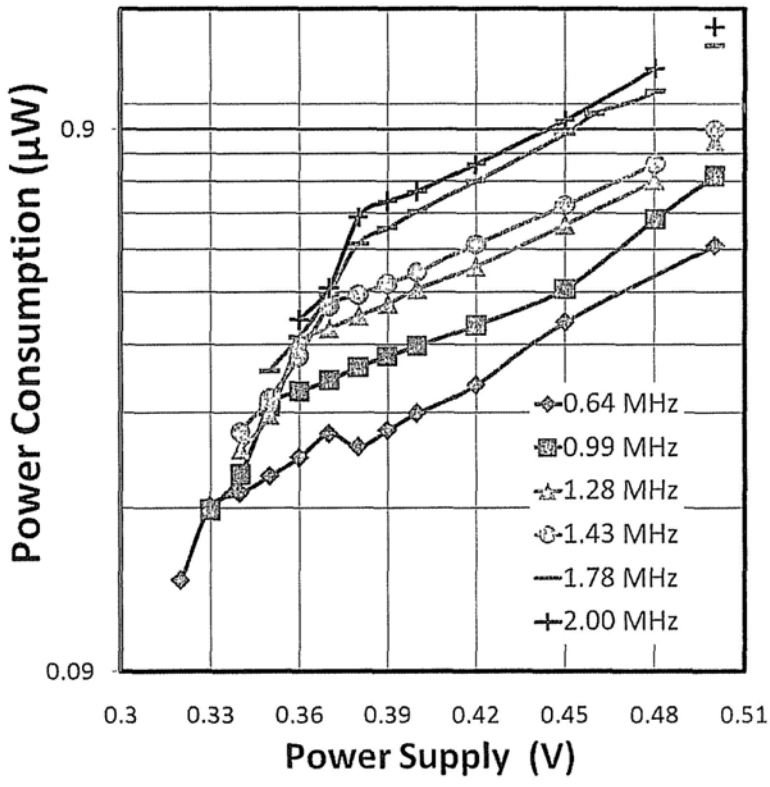


Figure 6.11 Power consumption of the BBP versus power supply with different input clock frequency, 90 nm technology

6.1.4 Chip Measurement: BBP Integrated in Tag

The BBP design of basic version is integrated in the entire RFID passive tag design and was fabricated in UMC 180nm, 130nm and 90 nm CMOS technology respectively. Figure 6.12 is the die photo of the entire tag design in 90 nm technology.

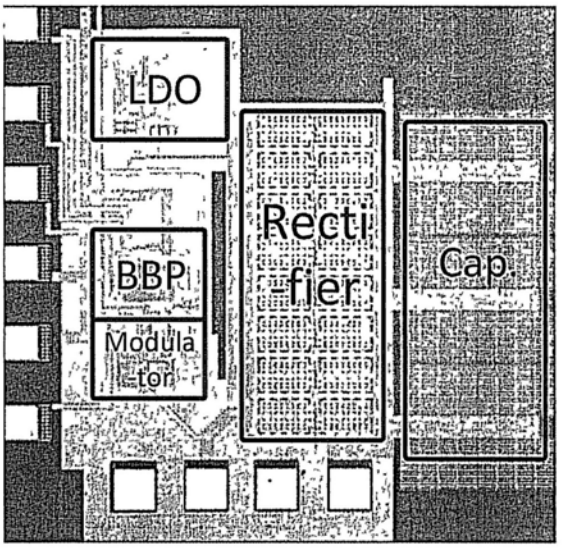


Figure 6.12 Die photograph of the entire tag in 90 nm CMOS technology

In the measurement, as shown in Figure 6.13(a), the FPGA board generates PIE encoded commands with iteration into the RF signal generator. And then the signal generator sends the RF signal modulated by the FPGA "reader", to the chip of tag. The tag is powered up this modulated RF signal and consequently the internal BBP starts to work. If the received command is valid and matches the condition of backscattering, the BBP will control the large NMOS device (the backscatter modulator) to reflect the response back. By applying the RF circulator, the oscilloscope can monitor the RF communication between the signal generator and the chip of tag without interference. In the measurements of the tags fabricated in the three technologies, the tag design can function at a minimum input power around -10 ~ -15 dBm. The tag fails to work due to other block's failure at a lower input power. The measured minimum power supply for the internal BBP designs are approximately 0.5 V, but in the individual test of these BBP designs, they can survive at a lower VDD. Figure 6.13(b) is the captured waveform in the measurement of the entire tag in 90 nm process.

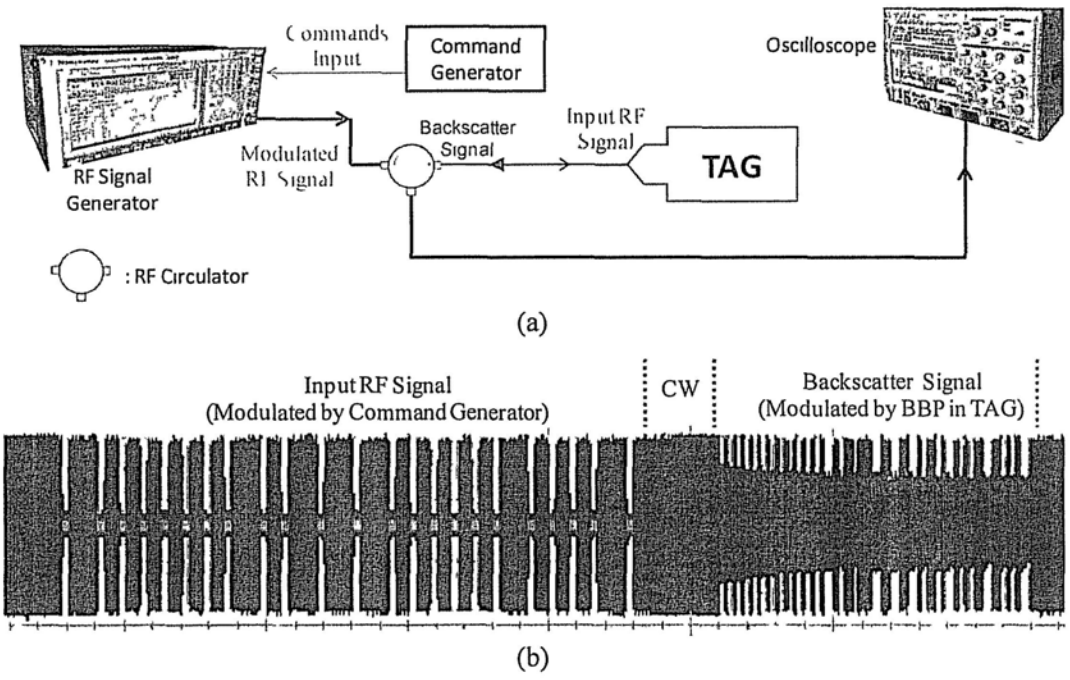


Figure 6.13 (a) Environment of the RFID tag measurement; (b) Captured waveform in the RFID tag measurement

6.1.5 Chip Measurement: Summary

From individual BBP design to the integrated basic version in the tag design, from the 180 nm CMOS technology to the more advanced 90 nm technology, the BBP design for subthreshold operation has been verified in different aspect for the robustness of the structure. Table 6.1 lists the chip specification and measured results of all the fabrications.

Table 6.1 Chips Measurement Summary

Technology	180 nm	130 nm	90 nm
Designs on chips	Individual BBP (basic version), Tag design with internal BBP (basic version)	Individual BBP (complex & basic version), Tag design with internal BBP (basic version)	Individual BBP (complex & basic version), Tag design with internal BBP (basic version)
Regular Threshold Voltage	0.52 V	0.45 V	0.41 V
Operating Voltage	BBP (basic version): 0.4~1.2 V	BBP (complex version): 0.33~0.9 V PIE Decoder (in the complex BBP): 1.05 ~ 0.3 V BBP (basic version): 0.37 ~ 1.0 V	BBP (complex version): 0.33~0.7 V BBP (basic version): 0.35 V at minimum
Minimum Full-function Power Consumption*	BBP (basic version): 180 nW	BBP (basic version): 136 nW BBP (complex version): 238 nW	BBP (basic version): 100 nW BBP (complex version): 179 nW

* The average power consumption is partially dependent on the data rate and the input command frequency specified by the reader, here the minimum power is measured at the minimum power supply and minimum input clock frequency when the *Query* command is input to the BBP in iteration

There are few publications reported in recent years for low power, low voltage RFID baseband processor design. And the processes for fabrication are 180 nm CMOS technology and the previous technology nodes. In comparison, the proposed designs show advanced performance in minimum operating voltage, as shown in Table 6.2. In power consumption comparison, as mentioned in Table 6.1, the power consumption is dependent on several variables. Specifically, besides the operating voltage, the data rate and how frequently the commands are sent to the BBP, what kinds of commands are sent, define the power dissipation in a certain period. Therefore the values in Table 6.2 are just reported minimum power consumption in the published works. As these issues are not all available from the publications, quantity power comparison is lack of necessary conditions. In general, the proposed design has lower power than these published works as it has a much lower supply voltage.

Table 6.2 Comparison of BBP designs for RFID passive tag

Design & Process	Protocol of Standard	Minimum Fastest Clock	Minimum Operating Voltage	Reported Minimum Power*
From [49], 180 nm	EPC C1G2	1.98 MHz	1 V	6400 nW
From [50], 180 nm	EPC C1G2	1.98 MHz	0.6 V	800 nW
From [51], 180 nm	ISO 18000-6B	0.8 MHz	0.6 V	440 nW
This work (basic version) 180 nm	EPC C1G2	1.98 MHz	0.4	180 nW
This work (complex version) 130 nm	EPC C1G2	1.98 MHz	0.33 V	238 nW

* There is no standard critera on how to define or evaluate the minimum power, here the minimum power is just the reported value from the publications

6.2 ACRL Circuits Verification

The fabricated ACRL cells, complex circuits, and related test circuits have been described in detail in Chapter 5. Figure 6.14 is the die photo of them, which are implemented in the target process: 130 nm CMOS technology.

6.2.1 ACRL Cells in Chains & Ring

In the test of 2-input XOR rings, the power supply scales down from 0.35 V to 0.15 V, the oscilloscope monitor and measure the oscillating waveform from the *selected Out* PAD in Figure 5.21.

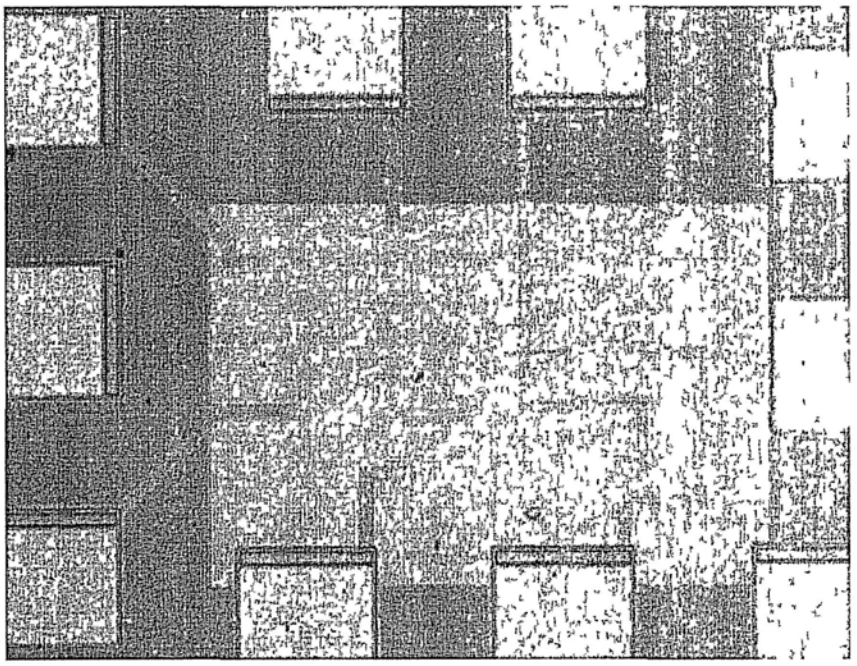


Figure 6.14 Die photograph of the ACRL circuits and test circuits

To further inspect the ACRL cell's superiority of speed, the 2-input XOR gate in dynamic CMOS style (Figure 6.15 (a)) is also applied in the layout of a 20-stage XOR chain, to compare with the ACRL XOR gate Charge sharing problem may cause logic error if input A and B change from HIGH to LOW in the evaluation period as node X should be keep at HIGH. For other AOI-like logic function with more inputs, this problem can be more serious.

Another logic-error problem is the pull-down leakage at node X in long period of evaluation. Especially when the subthreshold power supply scales down lower, node X tends to drop below the 50% of V_{DD} in a comparatively shorter interval during the period of evaluation. Figure 6.15 (b) is the interval that node X can hold above 50% V_{DD} versus the power supply V_{DD} in the period of evaluation Extra clock insertion and consideration on netlist design are also required if dynamic logic is adopted. Additionally, the footed NMOS and the output buffer make the dynamic XOR gate a longer delay (see Figure 6.16). Therefore ACRL logic should be comparatively a better choice in subthreshold application.

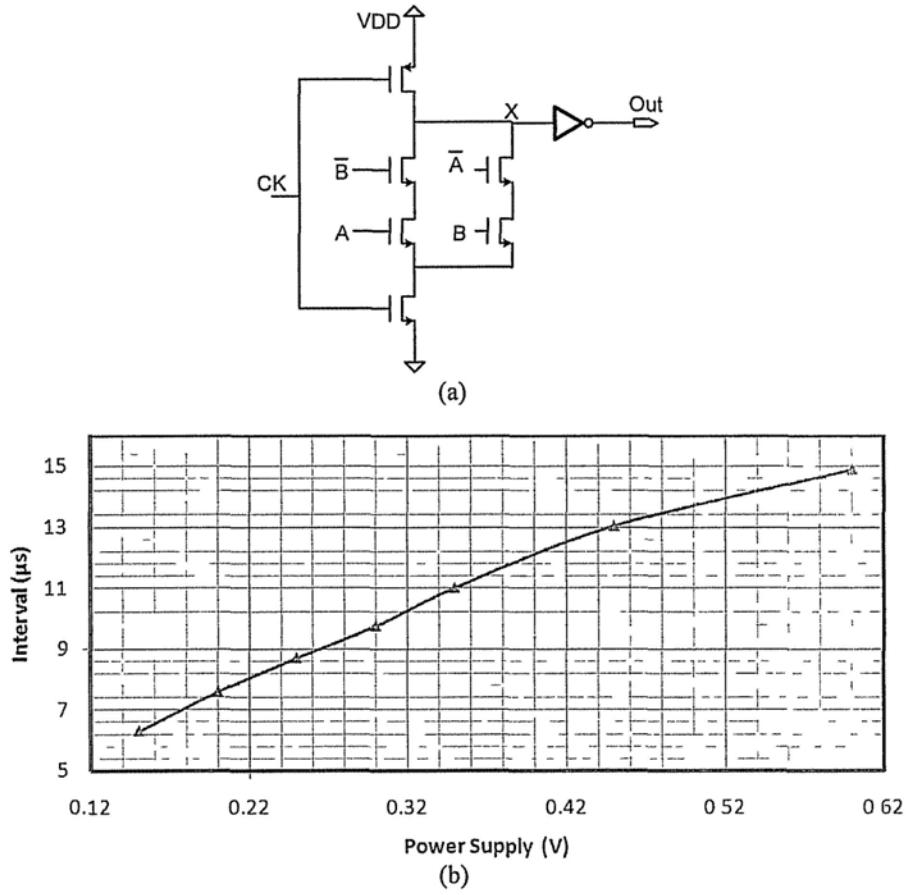


Figure 6.15 (a)Dynamic 2-input XOR gate circuit; (b) the time interval that node X can hold at HIGH (50%V_{DD}) in evaluation period

Derived the average delay time by the Equation (5.3), the measured and simulated delay time of the 2-input XOR gate rings are plotted in Figure 6.16 versus the supply voltage. In the comparison of measured results, the measure delay of ACRL cell results in a 40~60% reduction when compared to the measured standard cell. In the comparison of simulated and measured results, there is 20~50% error between the subthreshold simulation and the physical measurement. This is due to the lack of high accuracy in subthreshold modeling in the library. In the comparison of ACRL and dynamic CMOS XOR gate, the ACRL logic has a 25~30% delay reduction. Figure 6.17 is the oscillating waveform of the ACRL XOR ring in measurement.

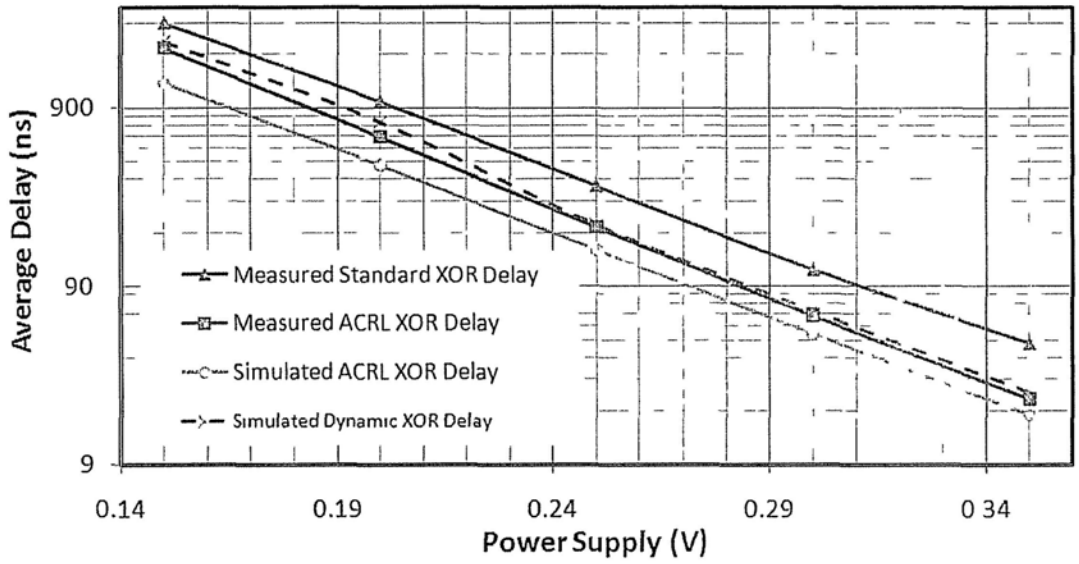


Figure 6.16 Measured and simulated delay time of 2-input XOR gates

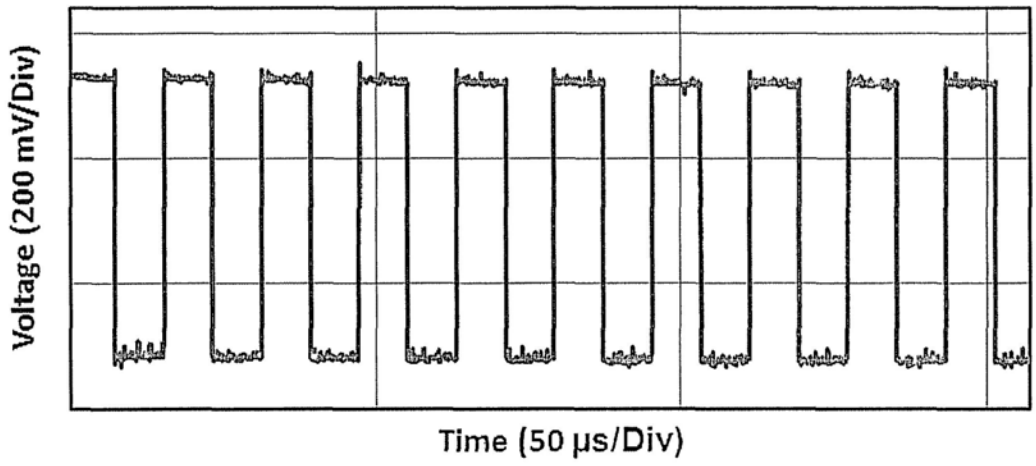


Figure 6.17 Oscillation of the ACRL XOR ring in measurement

To test the 2-bit equality comparator and the 4-bit GT comparator, a slow clock is applied in the *Control IN* PAD in Figure 5.21, and then measure the total delay: the timing difference of the voltage transitions between the *Selected Out* PAD and the *Dummy Out* PAD. After being divided by 16, the average delay of the cell in the 16-stage chain is calculated out. Figure 6.18 and 6.19 are the measured average delays comparison of the 2-bit equality comparator and the 4-bit GT comparator versus VDD respectively.

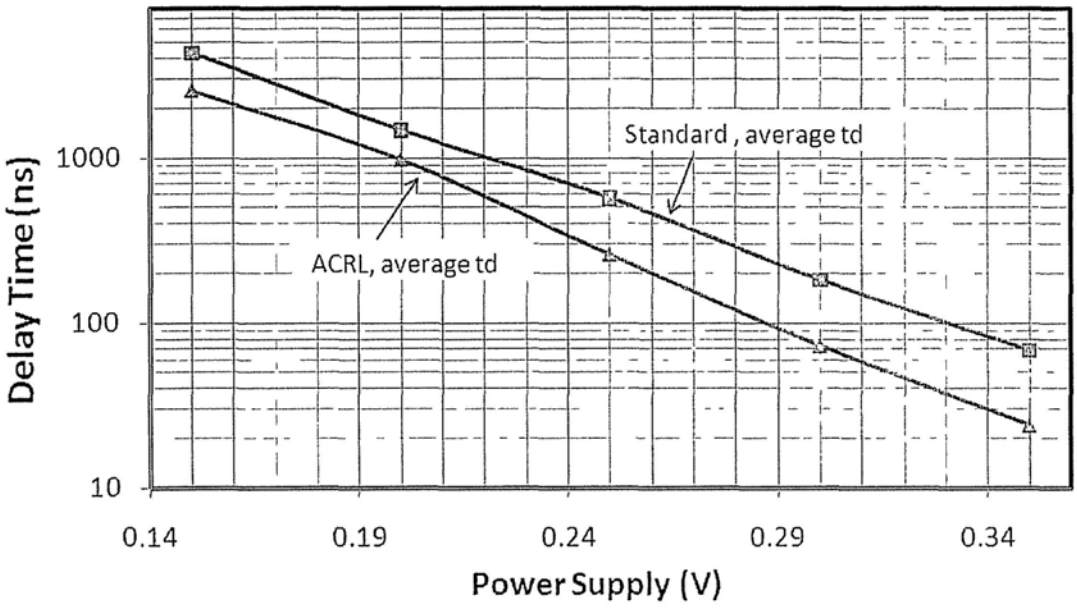


Figure 6.18 Measured average delay of the 2-bit equality comparator

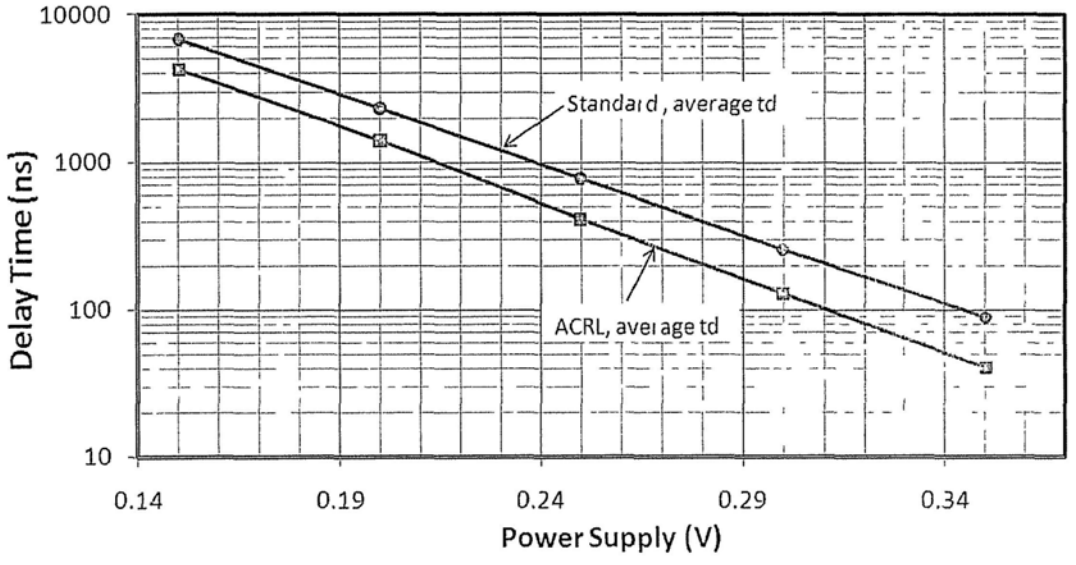


Figure 6.19 Measured average delay of the 4-bit GT comparator

In the measurement of the rising-edge-triggered DFF chains, clock signal is also applied to the Control IN PAD. When the 8-stage DFF chain is clocked to the $\bar{Q}$ outputs as "10000000", the next rising edge will triggered a propagation of 7 CK-to-$\bar{Q}$ rising delay and one CK-to-$\bar{Q}$ falling. When the 8-stage DFF chain is clocked to the $\bar{Q}$ outputs as "00000000", the next rising edge will triggered a propagation of 8 CK-to-$\bar{Q}$ rising delay. In this way the CK-to-$\bar{Q}$ rising and falling delays of a single

DFF cell can be derived. Figure 6.20 is the delay time comparison of the single ACRL DFF and standard DFF, versus power supply.

The power consumption is not measured in ACRL circuits as they share the same power supply and hard to be tested. In comparison to the power consumption of standard cells, ACRL cells consume static power when the output is at LOW. This is due to the pull-up and pull-down strength fighting. As the power supply scales down to the subthreshold region, the magnitude of leakage current is more and more close to the static current, and the time span of the dynamic current becomes longer in the logic transition of standard cells. Given a high logic activity α, the ACRL cell's power consumption can be close to the standard cell's one, and even lower than that. Figure 6.21 is the post-layout simulation of the current consumption of ACRL and standard 4-bit equality comparators. Here we use input signal's frequency to denote the logic activity. And it can be observed that as the frequency goes higher (higher logic activity), the power consumption of the ACRL cell increase much slower and becomes lower than the standards cell's power.

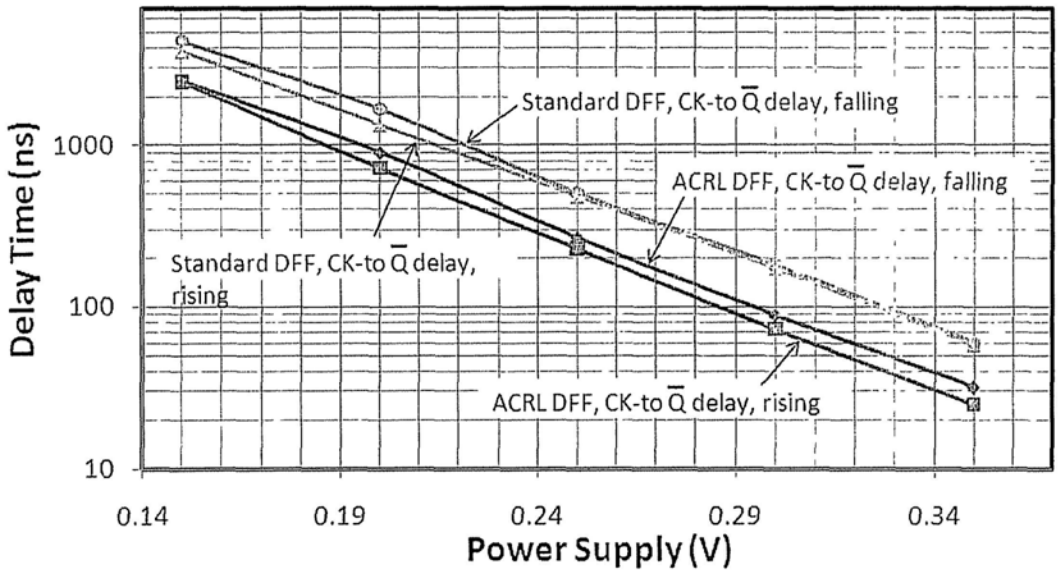
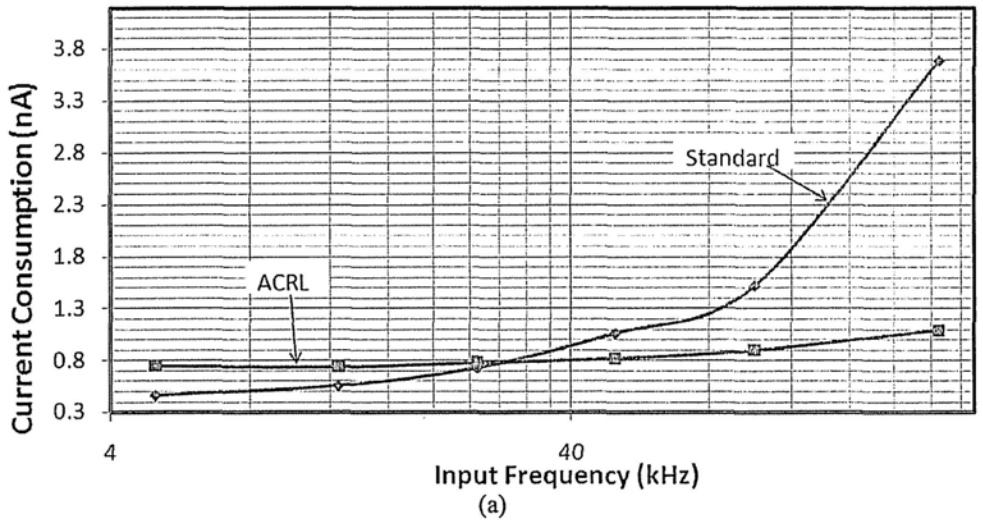
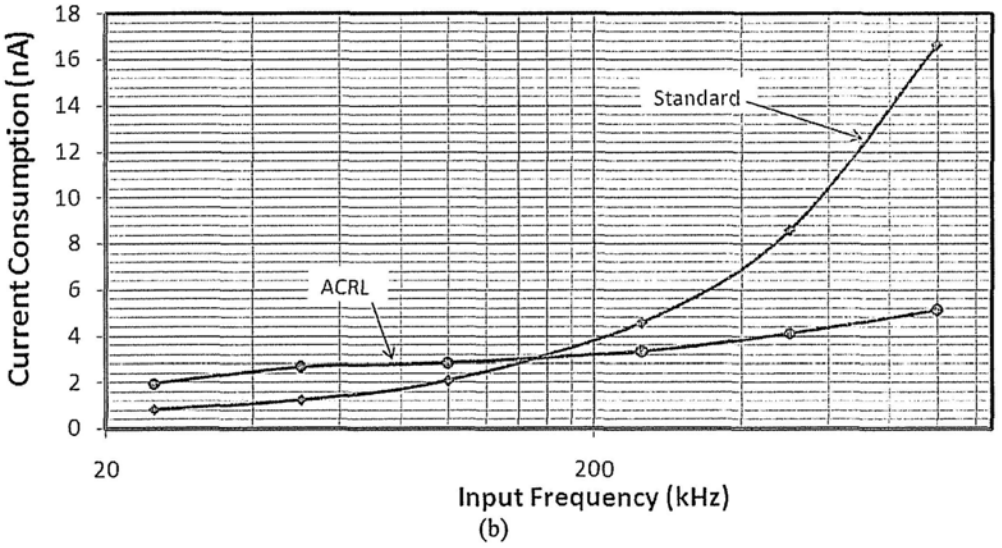


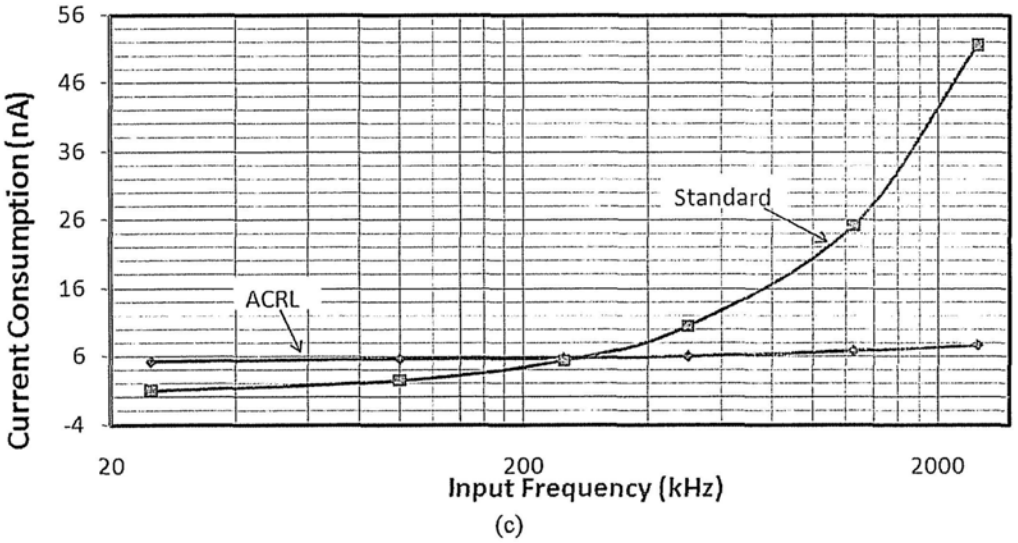
Figure 6.20 Measured CK-to-$\overline{Q}$ delay of the DFF



(a)



(b)



(c)

Figure 6.21 Current consumption of ACRL and standard 4-bit equality comparator versus logic activity at (a) 0.2 V; (b) 0.25 V; (c) 0.3 V power supply

6.2.2 Complex ACRL Circuits

In the testing of the 32-bit equality comparator, the test pattern (the 32-bit input of the comparator) and the select signal are serially shifted in to the 35-bit chain DFF in Figure 5.22. The measurement is carried out as instructed in sub Chapter 5.3.2. Some test input patterns are selected to verify the function and performance of the comparators. Figure 6.22 gives the measurement results of delays in comparison. The delay time reduction is evident, and the amount can be over 1 μs as the voltage scales down.

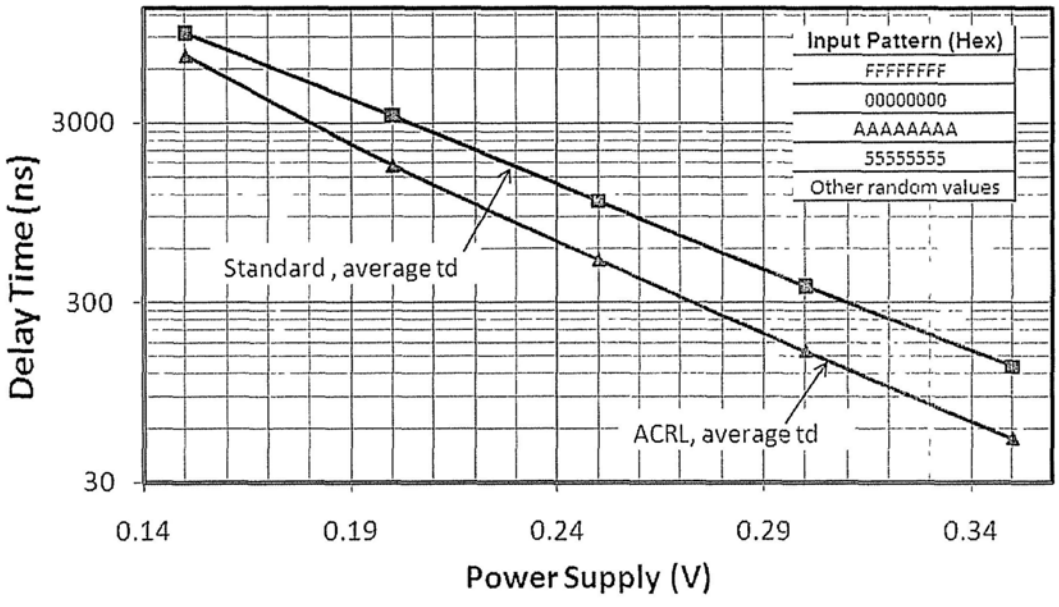


Figure 6.22 Measured average delay of 32-bit equality comparator versus VDD

In the test of the ACRL PIE decoder, the FPGA board generates the PIE formatted command, and then input to the *Control IN* PAD after voltage division. The decoded data comes out from the *Data Out* PAD and it is monitored and measured by the oscilloscope. The ACRL PIE decoder can function normally with all the EPC C1G2 specified data rates with minimum power supply 0.26 V, a 40 mV improvement to the standard PIE decoder design proposed in Chapter 4.

Figure 6.23 is the samples average delay from the rising edge of the PIE input to the output of the decoded data. Because different internal signals' values can cause different delays, the samples in measurement can't cover all the conditions, therefore Figure 6.23 is just a roughly profile of the PIE decoding delay versus power supply. Figure 6.24 is the captured waveform of the PIE command input and the output of the decoded data in measurement.

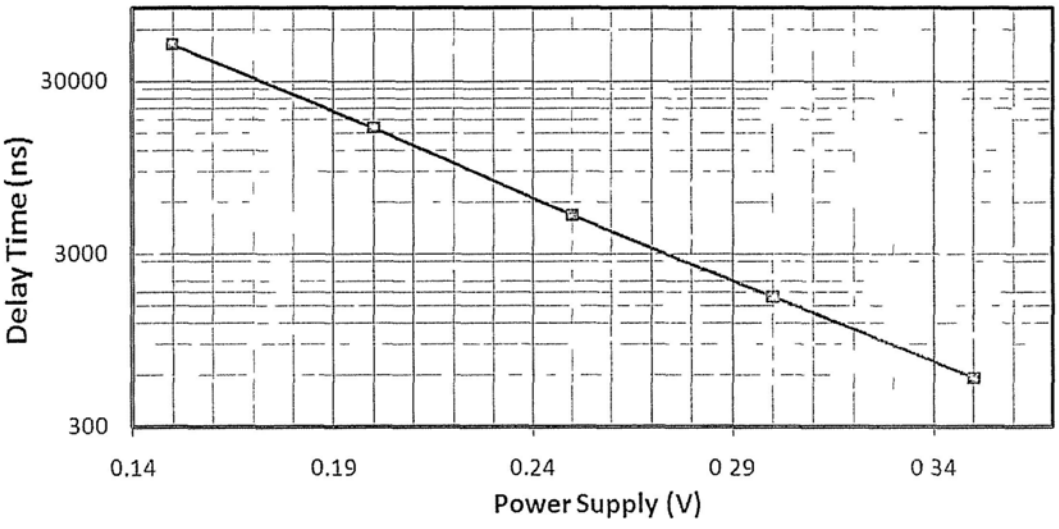


Figure 6.23 Average delay of the ACRL PIE decoder in measurement

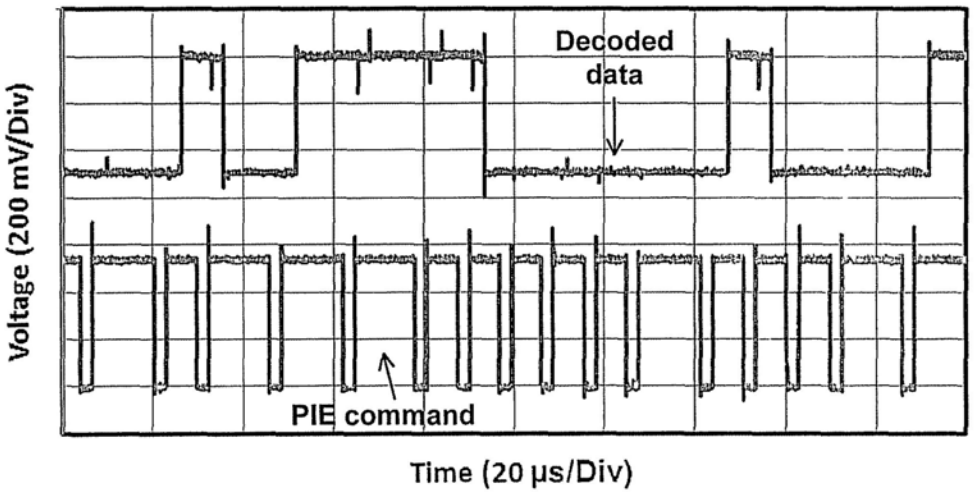


Figure 6.24 Captured waveform in the measurement of the ACRL PIE decoder

6.2.3 Summary

The ACRL cells and complex circuits' measurement have verified the subthreshold operation of the proposed ACRL designs in Chapter 5. The performance of the ACRL topology results in 30~70% reduction of the logical delay when compared to the standard cell designs. In the aspect of power consumption, the ACRL cells exhibits close or lower power in subthreshold high-logic-activity operation. In the area comparison, the custom layout of ACRL cells exhibits a slight smaller area occupation when compared to the standard static CMOS cells. Table 6.3 shows some area comparisons of these two logic styles.

Table 6.3 Comparison of area of ACRL and standard CMOS cells

Name	**2-input XOR**		**4-bit EQ Comparator**		**32-bit EQ Comparator**		**DFF**	
	Stand -ard	ACRL	Stand -ard	ACRL	Stand -ard	ACRL	Stand -ard	ACRL
Area (μm^2)	21.42	16.46	105.98	78.21	875.21	691.68	28.736	33.48

CHAPTER 7 CONCLUSION & FUTURE WORK

7.1 Conclusion

In this thesis hierarchical study and design have been carried out for a better understanding and improvement in digital circuit's subthreshold operation. The EPC C1G2 baseband processor for passive UHF RFID tag, a small digital system, has been selected as an example of subthreshold low power design. In the system level, the BBP is partitioned according to the clock domains and the requirement of subtrhshold operation. In the detailed RTL level design, innovative structures are proposed to relax the subthreshold timing. Pipelining, parallelism, clock gating and one-hot state transition are implemented in the logic design according to the actual requirement. The corresponding analysis on these improvements indicates a better robustness in subthreshold operation. In the transistor level design, based on the example design and general needs in digital circuits, a novel logic style is proposed to speed up the subthreshold logical signal propagation. Belong to the category of ratioed circuits, the proposed active control ratioed logic (ACRL) style is derived from previous logic styles but with obvious modification. It is a logic style with fast pull-up network and less capacitance, especially suitable for the implementation of high fan-in AOI structure. With related configuration the ACRL circuit can achieve a wide range of logic functions.

The numerical analysis and circuit simulations support the feasibility of the proposed designs in different levels. And in the physical verification, the chip measurements have confirmed the robustness and performance of them. The proposed logic architecture helps the baseband processor to function normally at a power supply 100 mV below the regular threshold voltage of the target process.

Fabricated in other deep submicron CMOS technologies and with robust functionality at subthreshold-voltage testing, the proposed logic design also demonstrates that it is feasible in multiple processes. The performance of the measured ACRL circuits result in 1.4X~3X logic propagating speed when compared to the standard cell designs. With the custom ACRL cells design, the proposed complex ACRL circuits also exhibit better performance in measurement. As long as the working frequency or logic activity is high enough, the ACRL circuits can have close to or lower power consumption in subthreshold operation, in comparison to standard-cell circuits.

Besides the proposed practical designs, this thesis also provides corresponding methodologies and studies on subthreshold digital design. Together with the detailed description and result of the designs, the thesis is a suitable reference for very-low-voltage and low power digital design.

7.2 Future Plan

The ARCL cells and circuits have been designed and verified, but it is still a long way to the full characterization and applicable to digital cell library. Good control must be excercised as the ACRL cells suffer from static power consumption in idle state or low active period. Therefore, in future work, there are several items on the waiting list:

(1) Redraw or modify the ACRL cells layouts to formal style that satisfies the digital cell library requirements.

(2) Try more fabrication runs and testing, and utilize the EDA tools to characterize accurate timing matrix according to post-layout simulation and measured results.

(3) Embed characterized ACRL cells into digital cell library and implement an ACRL digital design with auto placement and routing

(4) Try to develop more cell circuits suitable to the ACRL structure.

(5) Try to develop power saving scheme by controlling the NMOS loads at the pre-drive stage of ACRL structure.

REFERENCE

[1] A. Wang, Benton H. Calhoun, Anantha P. Chandrakasan, "Sub-threshold Design for Ultra Low-Power Systems," Springer Science +Business Media, LLC, ISBN-10: 0-387-33515-3, 2006.

[2] Armin Tajalli, Elizabeth J. Brauer, Yusuf Leblebici, and Eric Vittoz, "Subthreshold Source-Coupled Logic Circuits for Ultra-Low-Power Applications," IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 43, NO. 7, pp. 1699-1710, JULY 2008

[3] R. Roy, R. Krishnammthy, "Design of Low voltage CMOS circuits," The IEEE International Symposium on Circuits and Systems, Tutorial, May 6-9, 2001

[4] E. Vittoz' notebook, CEH, 1967

[5] SWANSON, MEINDL, "Ion-Implanted Complementary MOS Transistors in Low Voltage circuits," JSSC, VOL. SC-7,NO. 2, APRIL 1972

[6] Bipul C. Paul, Arijit Raychowdhury and Kaushik Roy, "Device Optimization for Digital Subthreshold Logic Operation," IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 52, NO. 2, pp. 237-247, FEBRUARY 2005

[7] Christian C. Enz and Eric A. Vittoz, "Charge-based MOS Transistor Modeling-The EKV model for low-power and RF IC design," John Wiley & Sons, Ltd, 2006

[8] Klaus von Arnirn, Eduardo Borinski and et. Al, "Efficiency of Body Biasing in 90 nm CMOS for Low Power Digital Circuits", IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 40, NO. 7, JULY 2005

[9] Norbert R Malik, "Electronic circuits: analysis, simulation, and design. Englewood Cliffs," NJ: Prentice Hall, pp. 315–316, 1995, ISBN 0-02-374910-5. http://worldcat.org/isbn/0-02-374910-5.

[10] S.M. Sze, "Semiconductor Devices: Physics and Technology," Wiley, 1985

[11] ANDREAS G. ANDREOU and KWABENA A. BOAHEN, "Translinear Circuits in Subthreshold MOS," Analog Integrated Circuits and Signal Processing, 9, pp. 141-166, 1996

[12] "ECE-E302, Electronic Devices, Lab III: Capacitance of a PN Junction," Drexel University

[13] Paul, Raychowdhury and K. Roy, "Device Optimization for Digital Subthreshold Logic Operation", IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 52, NO. 2, FEBRUARY 2005

[14] Hendrawan Soeleman, Kaushik Roy, "Digital CMOS Logic Operation in the Sub-Threshold Region," Proceedings of the 10th GLSVLSI, 2000

[15] B. H. Calhoun, A. Wang, A. Chandrakasan, "Modeling and Sizing for Minimum Energy Operation in Subthreshold Circuits," JSSC, VOL. 40, NO. 9, SEPTEMBER 2005

[16] B. H. Calhoun, A. Wang and et. al, "Sub-Threshold Design: The Challenges of Minimizing Circuit Energy," ISLPED'06

[17] A. Wang, A. Chandrakasan, "A 180-mV Subthreshold FFT Processor Using a Minimum Energy Design Methodology," JSSC, VOL. 40, NO. 1, JANUARY 2005

[18] John Crossley, Allen Lee, "Analysis of Design Techniques for Energy-Optimized SRAM," Dept. of EECS, Univ. of California, Berkeley

[19] Naveen Verma, Anantha P. Chandrakasan, "A 65nm 8T Sub-Vt SRAM Employing Sense-Amplifier Redundancy," SESSION 18, 18.4, ISSC 2007

[20] Hendrawan Soeleman, Kaushik Roy, Bipul C. Paul, "Robust Subthreshold Logic for Ultra-Low Power Operation", IEEE TRANSACTIONS ON VERY

LARGE SCALE INTEGRATION (VLSI) SYSTEMS, VOL. 9, NO. 1, FEBRUARY 2001

[21] Armin Tajalli, Elizabeth J. Brauer, "Subthreshold Source-Coupled Logic Circuits for Ultra-Low-Power Applications", JSSC, VOL. 43, NO. 7, JULY 2008

[22] Fariborz Assaderaghi, Dennis Sinitsky, et. al., "Dynamic Threshold-Voltage MOSFET (DTMOS) for Ultra-Low Voltage VLSI," IEEE TRANSACTIONS ON ELECTRON DEVICES, VOL. 44, NO. 3, MARCH 1997

[23] Tajalli, A., Vittoz, E., et. al., "Ultra Low Power Subthreshold MOS Current Mode Logic Circuits Using a Novel Load Device Concept," Solid State Circuits Conference, 2007, Sep., 2007

[24] Chris Hyung-Il Kim, Hendrawan Soeleman, and K.Roy, "Ultra-Low-Power DLMS Adaptive Filter for Hearing Aid Applications", IEEE TRANS.S ON VLSI SYSTEMS, VOL. 11, NO. 6, DECEMBER 2003

[25] Benton H. Calhoun, Anantha P. Chandrakasan, "A 256kb Sub-threshold SRAM in 65nm CMOS," IEEE International Solid State Circuits Conference (ISSCC), pp. 628-629, February 2006.

[26] Ik Joon Chang, Jae-Joon Kim, Sang Phill Park, K. Roy, "A 32kb IOT Subthreshold SRAM Array with Bit-Interleaving and Differential Read Scheme in 90nm CMOS", ISSCC, 2009

[27] Benton H. Calhoun, Anantha P. Chandrakasan, "Static Noise Margin Variation for Sub-threshold SRAM in 65-nm CMOS", JSSC, VOL. 41, NO. 7, JULY 2006

[28] Tae-Hyoung Kim, Jason Liu, John Keane, Chris H. Kim, "A High-Density Subthreshold SRAM with Data-Independent Bitline Leakage and Virtual

Ground Replica Scheme," IEEE International Solid State Circuits Conference (ISSCC), pp. 330-331, February 2007.

[29] Jaydeep P. Kulkarni, Keejong Kim, and Kaushik Roy, "A 160 mV Robust Schmitt Trigger Based Subthreshold SRAM," IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 42, NO. 10, OCTOBER 2007

[30] Mika Kontiala, Mika Kuulusa, and Jari Nurmi, " COMPARISON OF STATIC LOGIC STYLES FOR LOW-VOLTAGE DIGITAL DESIGN", Electronics, Circuits and Systems, 2001. ICECS 2001

[31] Fang-shi Lai and Wei Hwang, " Design and Implementation of Differential Cascode Voltage Switch with Pass-Gate (DCVSPG) Logic for High-Performance Digital Systems", JSSC, VOL. 32, NO. 4, APRIL 1997

[32] TI-RFID, "UHF Gen 2 System Overview," Texas Instruments, Mar., 2005

[33] M. Roberti, "A 5-cent breakthrough," http://www.rfidjournal.com.

[34] T. Umeda et. al., "A 950-MHz Rectifier Circuit for Sensor Network Tags With 10-m Distance," IEEE J. Solid-State Circuits, vol. 41, pp. 35-41, Jan.2006.

[35] Pavel V. Nikitin and K. V. S. Rao, "Performance Limitations of Passive UHF RFID Systems," Antennas and Propagation Society International Symposium 2006, IEEE

[36] Karthik Moncombu Ramakrishnan, Daniel D. Deavours, "Performance benchmarks for passive UHF RFID tags," 13th GI/ITG Conference on Measurement, Modeling, and Evaluation of Computer and Communication Systems, 2006

[37] D. Kim, M. A. Ingram, and W. W. Smith, "Measurements of small-scale fading and path loss for long range RF tags", IEEE Transactions on Antennas and Propagation, vol. 51, no. 8, pp. 1740-1749, Nov. 2003

[38] P. V. Nikitin and K. V. S. Rao, "Measurement of backscattering from RFID tags", Proceedings of AMTA Symposium, Oct. 2005

[39] Leena Ukkonen, Lauri Sydänheimo, "Performance Characterization of Passive UHF RFID Tags," The Internet of Things, 20th Tyrrhenian Workshop on Digital Communications, 2010

[40] Ka Nang Leung, Chiu Sing Choy, Kong-Pang Pun, et. al., "RF Module Design of Passive UHF RFID Tag Implemented in CMOS 90-nm Technology," J. Low Power Electron., VOL. 6, pp. 141-149, 2010

[41] Chiu Sing Choy, "Study the Design Challenges of 90nm Technology UHF RFID Tag IC," ITF Hong Kong, (ITP/017/07LP)

[42] TI-RFID, "TI Gen 2 UHF Protocol Reference Guide," Texas Instruments, June 2006

[43] EPCglobal Inc™, "EPC™ Radio-Frequency Identity Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860 MHz – 960 MHz Version 1.0.9," Jan. 2005.

[44] Ricci, A. et. al., "Design of a 2 μW RFID baseband processor featuring an AES cryptography primitive," Electronics, Circuits and System, ICECS 2008. 15th IEEE Intl. Conf. on, pp. 376-379, Sep. 2008.

[45] Impinj RFID Technology Series, "Gen 2 Tag Clock Rate - What You Need to Know," http://www.impinj.com

[46] Chi Fat Chan, Weiwei Shi, et. al., "A Low-Power Signal Processing Front-End and Decoder for UHF Passive RFID Transponders," Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on, pp.1581-1584, May. 2009.

[47] Neil H. E. Weste and David Harris, "CMOS VLSI Design: A Circuits and Systems Perspective 3rd Edition," Pearson Education Inc., 2005

[48] T. Simon, "A fast static CMOS NOR gate," Proc. Advanced Research in VLSI andParallel Systems, 1992, pp. 180-192.

[49] V. Roostaie, et. al., "A low power baseband processor for a dual mode UHF EPC Gen 2 RFID tag," Design and Tech. of Integrated Sys. in Nanoscale Era, 2008. 3rd International Conf. on, pp. 1-5, Mar. 2008.

[50] Ricci, A. et. al., "Design of a 2 μW RFID baseband processor featuring an AES cryptography primitive," Electronics, Circuits and System, ICECS 2008. 15th IEEE Intl. Conf. on, pp. 376-379, Sep. 2008.

[51] Andrea Ricci, Matteo Grisanti, Ilaria De Munari and Paolo Ciampolini, "Improved Pervasive Sensing With RFID: An Ultra-Low Power Baseband Processor for UHF Tags," IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATION (VLSI) SYSTEMS, VOL. 17, NO. 12, DECEMBER 2009

[52] Koichi Nose, Takayasu Sakurai, "Optimization of VDD and VTH for low-power and high speed applications," Proc. of the 2000 Asia and South Pacific Design Automation Conf., pp. 469-474, 2000.

[53] Z. Lao, J. Jensen, K. Guinn and M. Sokolich, "1.3 V supply voltage 38 GHz static frequency divider," ELECTRONICS LETTERS, 4th, Vol. 40, No. 5, March 2004

[54] Sung-Chuan Fang, Jyh-Ming Wang, and Wu-Shiung Feng, "A New Direct Design for Three-Input XOR Function on the Transistor Level," IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-I FUNDAMENTAL THEORY AND APPLICATIONS, VOL. 43, NO. 4, APRIL 1996

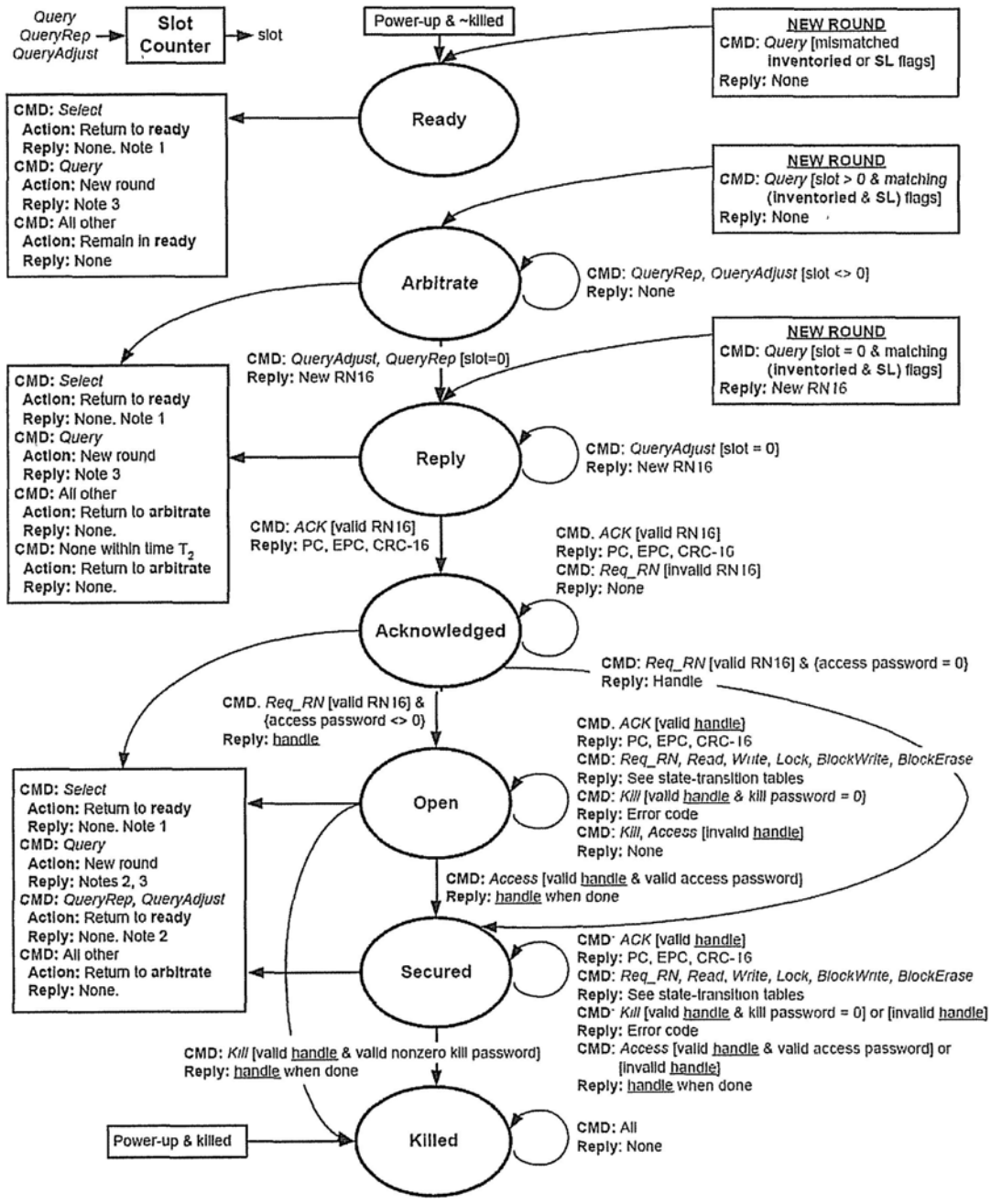
[55] Mohamed W. Allam and Mohamed I. Elmasry, "Dynamic Current Mode Logic (DyCML): A New Low-Power High-Performance Logic Style," IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 36, NO. 3, MARCH 2001

[56] Jose A. Rodriguez-Rodriguez, Manuel Delgado-Restituto and Angel Rodriguez-Vazquez, "Baseband-Processor for a Passive UHF RFID Transponder," International Conference on Green Circuits and Systems (ICGCS), pp. 344-348, June, 2010

[57] Von Arnim K., Seegebrecht P., Thewes, R. and Pacha, C., "A low-leakage 2.5GHz skewed CMOS 32b adder for nanometer CMOS technologies," IEEE International Solid-State Circuits Conference, Digest of Technical Papers, 2005

APPENDIX A

State Transition Diagram of EPC C1G2 Protocol



APPENDIX B

Description of commands (more details are in reference [43])

Select command

	Command	Target	Action	MemBank	Pointer	Length	Mask	Truncate	CRC-16
# of bits	4	3	3	2	EBV	8	Variable	1	16
description	1010	000 Inventoried (S0) 001 Inventoried (S1) 010 Inventoried (S2) 011 Inventoried (S3) 100 SL 101 RFU 110 RFU 111 RFU	See Table 6.19	00 RFU 01 EPC 10 TID 11 User	Starting Mask address	Mask length (bits)	Mask value	0 Disable truncation 1 Enable truncation	

Tag response to Action parameter

Action	Matching	Non-Matching
000	assert SL or inventoried → A	deassert SL or inventoried → B
001	assert SL or inventoried → A	do nothing
010	do nothing	deassert SL or inventoried → B
011	negate SL or (A → B, B → A)	do nothing
100	deassert SL or inventoried → B	assert SL or inventoried → A
101	deassert SL or inventoried → B	do nothing
110	do nothing	assert SL or inventoried → A
111	do nothing	negate SL or (A → B, B → A)

Query command

	Command	DR	M	TRext	Sel	Session	Target	Q	CRC-5
# of bits	4	1	2	1	2	2	1	4	5
description	1000	0 DR=8 1 DR=64/3	00 M=1 01 M=2 10 M=4 11 M=8	0 No pilot tone 1 Use pilot tone	00 All 01 All 10 ~SL 11 SL	00 S0 01 S1 10 S2 11 S3	0 A 1 B	0–15	

Tag reply to a *Query* command

	Response
# of bits	16
description	RN16

QueryAdjust command

	Command	Session	UpDn
# of bits	4	2	3
description	1001	00 S0 01 S1 10 S2 11 S3	110 Q = Q + 1 000 No change to Q 011 Q = Q − 1

QueryRep command

	Command	Session
# of bits	2	2
description	00	00 S0 01 S1 10 S2 11 S3

Tag reply to a *QueryAdjust* command

	Response
# of bits	16
description	RN16

Tag reply to a *QueryRep* command

	Response
# of bits	16
description	RN16

ACK command

	Command	RN
# of bits	2	16
description	01	Echoed RN16 or handle

Tag reply to a successful *ACK* command

	Response
# of bits	21 to 528
description	{PC, EPC, CRC-16} OR {00000_2, truncated EPC, CRC-16}

NAK command

	Command
# of bits	8
description	11000000

Req_RN command

	Command	RN	CRC-16
# of bits	8	16	16
description	11000001	Prior RN16 or handle	

Tag reply to a *Req_RN* command

	RN	CRC-16
# of bits	16	16
description	New RN16 or handle	

Read command

	Command	MemBank	WordPtr	WordCount	RN	CRC-16
# of bits	8	2	EBV	8	16	16
description	11000010	00 Reserved 01 EPC 10 TID 11 User	Starting address pointer	Number of words to read	handle	

Tag reply to a successful *Read* command

	Header	Memory Words	RN	CRC-16
# of bits	1	Variable	16	16
description	0	Data	handle	

Write command

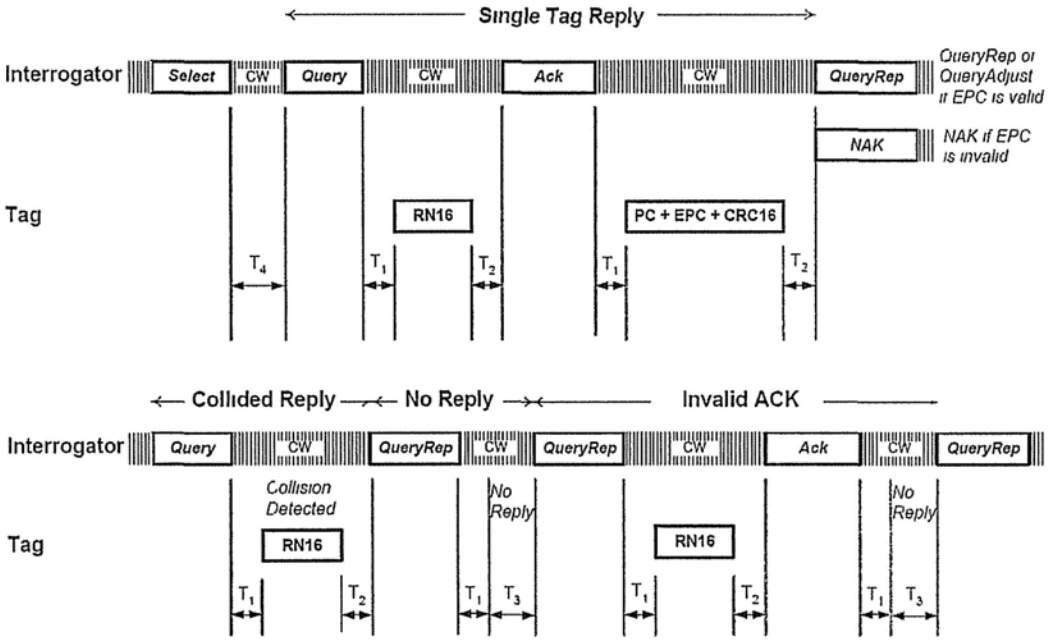
	Command	MemBank	WordPtr	Data	RN	CRC-16
# of bits	8	2	EBV	16	16	16
description	11000011	00 Reserved 01 EPC 10 TID 11 User	Address pointer	RN16 ⊗ word to be written	handle	

Tag reply to a successful *Write* command

	Header	RN	CRC-16
# of bits	1	16	16
description	0	handle	

APPENDIX C

Timing requirement of reader-tag communication (more details are in reference [43])



Link timing parameters

Parameter	Minimum	Typical	Maximum	Description
T_1	MAX(RTcal,10T_{pri}) × (1 − FT) − 2μs	MAX(RTcal,10T_{pri})	MAX(RTcal,10T_{pri}) × (1 + FT) + 2μs	Time from Interrogator transmission to Tag response (specifically the time from the last rising edge of the last bit of the Interrogator transmission to the first rising edge of the Tag response), measured at the Tag's antenna terminals
T_2	3 0T_{pri}		20 0T_{pri}	Time required if a Tag is to demodulate the Interrogator signal, measured from the last falling edge of the last bit of the Tag response to the first falling edge of the Interrogator transmission
T_3	0 0T_{pri}			Time an Interrogator waits, after T_1 before it issues another command
T_4	2 0 RTcal			Minimum time between Interrogator commands

Notes
1 T_{pri} denotes either the period of an FM0 symbol or the period of a single subcarrier cycle as appropriate
2 The maximum value for T_2 shall apply only to Tags in the reply or acknowledged states (see 6 3 2 4 3 and 6 3 2 4 4) For a Tag in the reply or acknowledged states if T_2 expires (i e reaches its maximum value)
- Without the Tag receiving a valid command the Tag shall transition to the arbitrate state (see 6 3 2 4 2)
- During the reception of a valid command the Tag shall execute the command
- During the reception of an invalid command the Tag shall transition to arbitrate upon determining that the command is invalid

In all other states the maximum value for T_2 shall be unrestricted "Invalid command is defined in 6 3 2 10
3 An Interrogator may transmit a new command prior to interval T_2 (i e during a Tag response) In this case the responding Tag is not required to demodulate or otherwise act on the new command and may undergo a power-on reset
4 FT is the frequency tolerance specified in Table 6 11
5 T_1+T_3 shall not be less than T_4

APPENDIX D

FM0 and Miller encoding (more details are in reference [43])

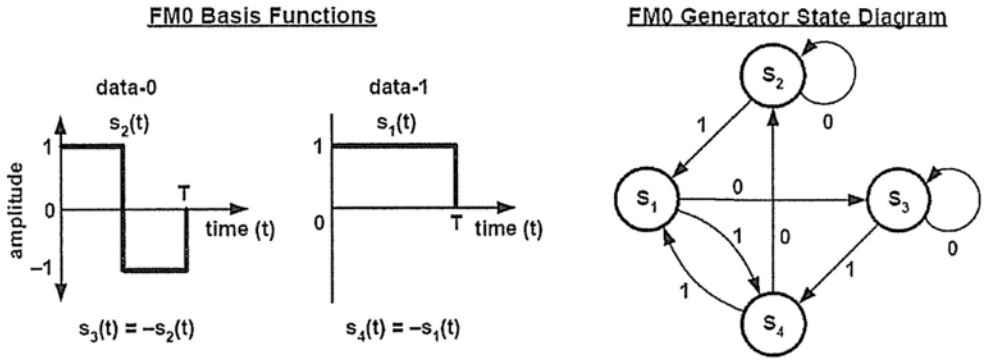
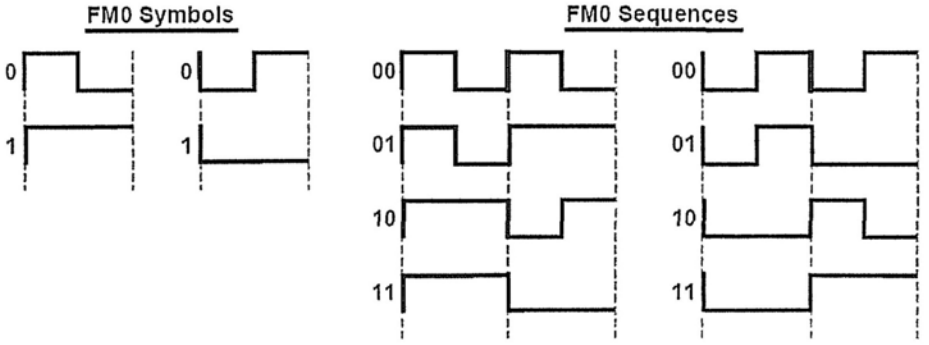
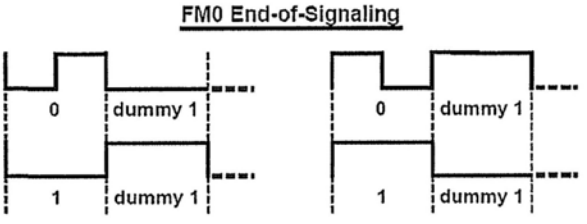


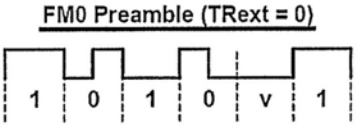
Figure 6.8 – FM0 basis functions and generator state diagram

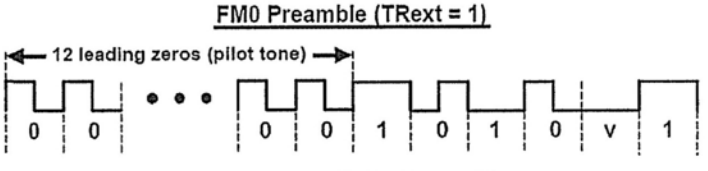


FM0 symbols and sequences

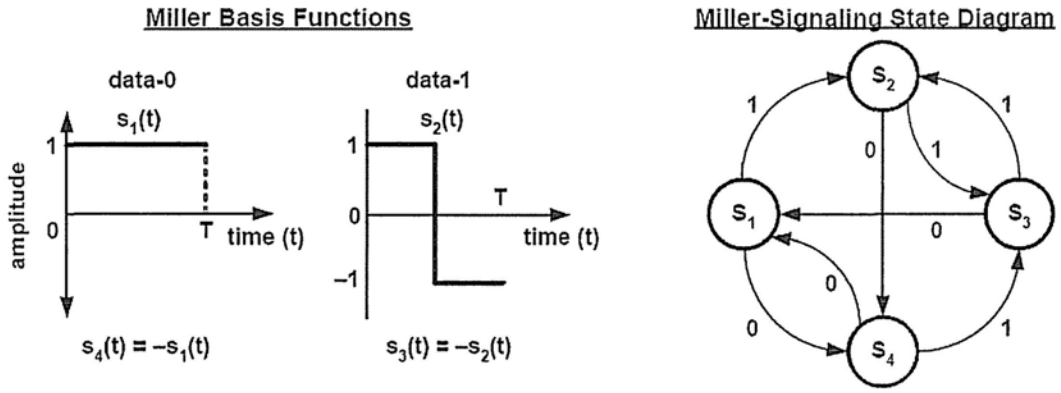


Terminating FM0 transmissions

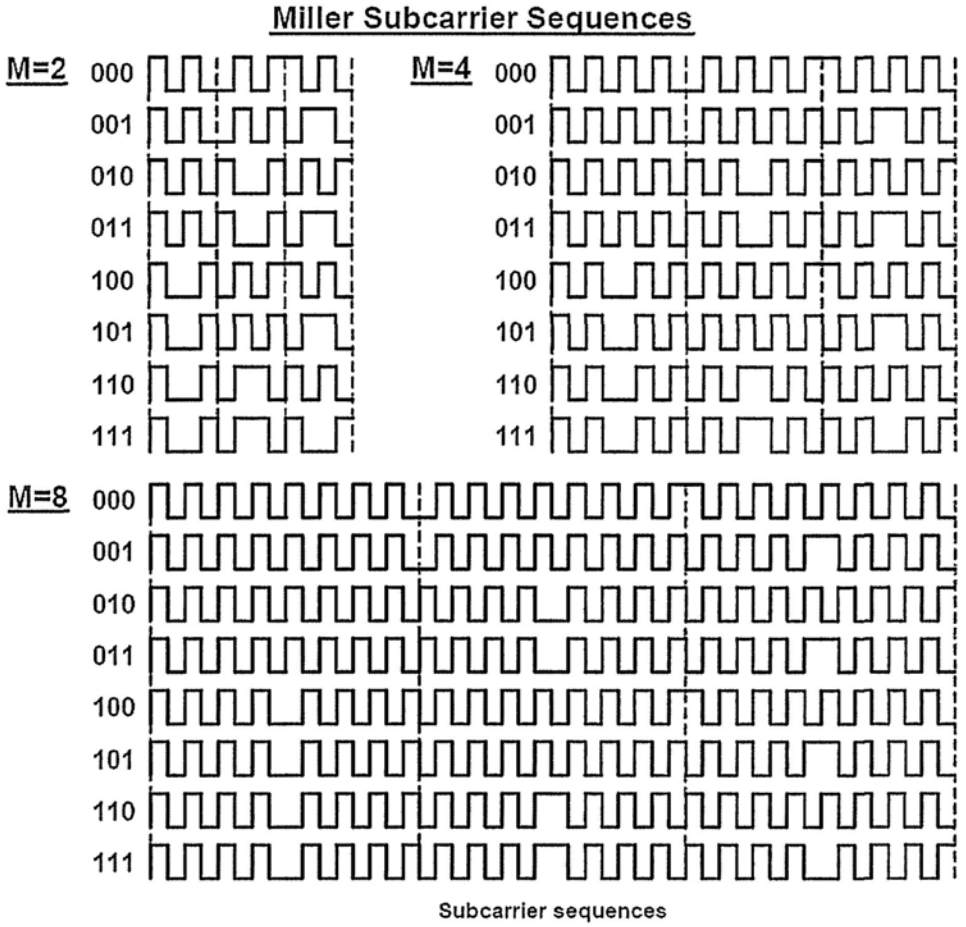




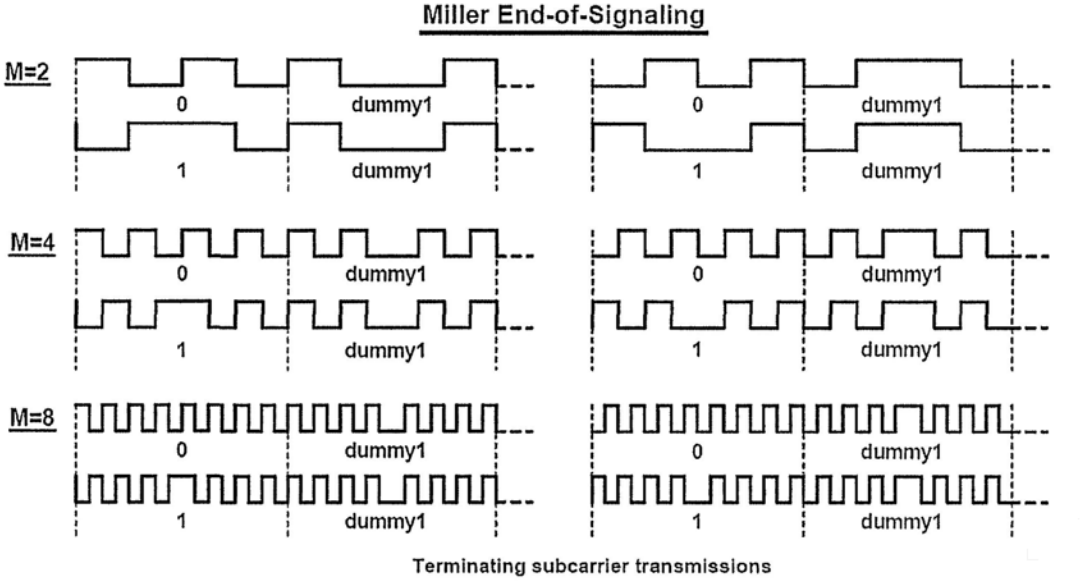
FM0 T=>R preamble



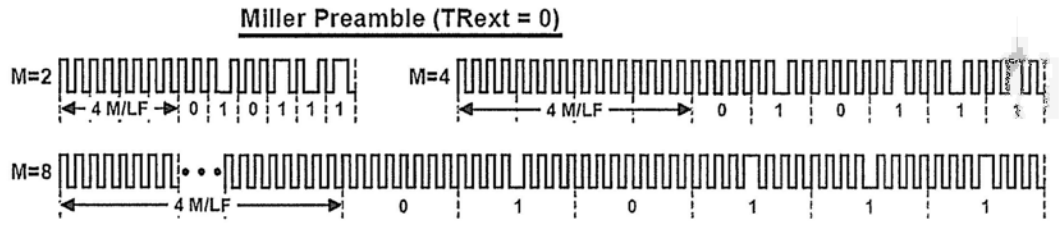
Miller basis functions and generator state diagram

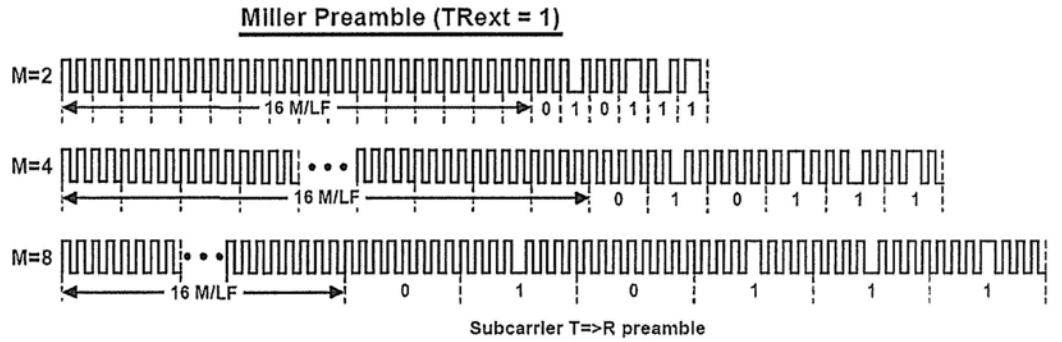


Subcarrier sequences



Terminating subcarrier transmissions





Subcarrier T=>R preamble