ROBUST ENERGY EFFICIENT DESIGN FOR ULTRA LOW VOLTAGE CMOS VLSI

D I S S E R T A T I O N

ZUR ERLANGUNG DES DOKTORGRADES

DER FAKULTÄT FÜR ANGEWANDTE WISSENSCHAFTEN
DER ALBERT-LUDWIGS-UNIVERSITÄT
FREIBURG IM BRESIGAU

von

(M.Sc) SENTHIL KUMAR JAYAPAL
REKTOR DER ALBERT-LUDWIGS-UNIVERSITÄT FREIBURG:
Prof. Dr. Hans-Jochen Schiewer

DEKAN FAKULTÄT FÜR ANGEWANDTE WISSENSCHAFTEN:
Prof. Dr. Hans Zappe

Gutacher:
1. Prof. Dr.-Ing. Yiannos Manoli
2. Prof. Dr. Doris Schmitt-Landsiedel

eingereicht: 03.07.2009
Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation mit dem Titel "ROBUST ENERGY EFFICIENT DESIGN FOR ULTRA LOW VOLTAGE CMOS VLSI" nur unter Verwendung der angegebenen Quellen und Hilfsmittel selbst angefertigt habe.

Freiburg, den 03.07.2009, – – – – – – – – – – – – – – – – – – – – – – – – – – – – – –

signature                                      (SENTHILKUMAR JAYAPAL)
this thesis is dedicated to my father, mother and my beloved sisters
Acknowledgements

The research work presented in this dissertation would not have been possible without kindness, support, motivation, guidance and inspiration from so many individuals. My research work as a doctoral student at the chair of microelectronics has been incredibly rewarding and engaging, and I wish to thank everyone who helped me to make it possible.

At first, I would like to thank my supervisor, Prof. Dr. (-Ing) Yiannos Manoli, for giving me an opportunity to work in his research group. Apart from his strong support and motivation, he encouraged me with kind supervision, guidance and showed me right direction throughout my research work. From his diversified research knowledge and expertise in both work and life gave me immense flexible, pleasure and comfortable during my research work.

I could not have been any luckier in terms of my fellow research group members. I would like to thank the members of our research group Dr.-Ing. Maurits Ortmanns, Dipl.-Phys. Joachim Becker and Dipl.-Ing. Markus Becker for their support, discussions and guidance. Next, I would like to thank our department secretary Elisabeth-Theresia Maurer and Claudia Jahn, who helped me to solve all my administration regulation issues. I consider myself very fortunate to have worked with such an amazing group of people. Finally, I would like to thank everyone in the group, for their support, encouragement and friendship. Further, I wish everyone in the chair of microelectronics a great success in their personal as well as in professional endeavors.

My family has been extremely helping me throughout this endeavor, and it is impossible to summarize how much they have helped in making this possible. So, I would like to thank my father, mother, elder and younger sister who given me the love and support to know that I can achieve anything that I can dream. Lastly, but certainly not least, I would like to thank my school and college friends for giving motivation and support both in my personal and professional life.
Abstract

Energy efficient circuit design becomes increasingly crucial for battery driven ubiquitous computing applications such as autonomous wireless sensor nodes, implantable bio-medical devices and other embedded portable gadgets. Due to the rapid scaling of process technology, the static leakage power becomes dominant over the active power consumption and both of which shorten the life time of the power source. To extend the battery life time while meeting the performance demands, the designers face critical challenges in choosing the appropriate logic circuit topologies with state-of-the-art leakage reduction techniques in order to optimally balance the power and performance trade-offs. Hence, we primarily focused on different sources of leakage current and effective techniques for minimizing the active and static leakage power consumption in the deep sub-micron regime.

To accomplish ultra-low power requirements, the sub-threshold circuit operation has emerged as a promising approach in the applications where the performance is not a primary concern. However, the higher impact of process variations and the low-performance solution minimizes the effectiveness of sub-threshold operation in the deep sub-micron process technology. To overcome these issues, we introduced the forward body bias in the ultra-low $V_{dd}$ region ($V_{dd} \leq 0.5V$) to determine the optimum between the energy and performance metric and further, to mitigate the influence of process variations. In addition, we studied the characteristics of symmetric and asymmetric forward body bias configurations in the sub-threshold, near-threshold and above-threshold region to determine the minimum energy consumption using an adder based delay chain with an industrial 130nm triple-well CMOS process technology. The desired optimum FBB of 0.4V and $V_{dd}/2$ is chosen based on several design parameters such as temperature, noise margin, device width, process variation and the circuit area overhead. Our result shows that, the
optimum minimum energy point exists in the sub-threshold region of 0.2V supply voltage with substantial performance improvement, and this minimum point and the respective performance advantages varies depending on the forward body bias configuration adapted. Furthermore, the symmetric FBB configuration reaches substantial reduction in the delay and active-energy variability compared to both the conventional zero body bias and 0.1V asymmetric forward body bias schemes.

Finally, the larger area overhead imposed by an external body bias circuitry and the increased active-leakage power from non-evaluation devices limits the advantage of conventional forward body bias approach in the ultra-low $V_{dd}$ regime. To utilize the advantage of forward body bias and to minimize the active-leakage power, we proposed a novel inherent self-adaptive body bias approach for the precharge-evaluate logic circuits. In this approach, the forward body bias is applied to high threshold voltage of either the pull-up or pull-down logic network by pre-conditioning the body terminal depending on the output evaluation transitions. This approach improves the energy efficiency with considerable performance gain at negligible area and power overhead, which is highly required for low-to-medium performance level applications. To validate the proposed method, the 16-bit carry look-ahead adder has been designed and simulated with static and dynamic Manchester carry chain for different body bias schemes. At skew ratio=1, the delay improvement of 15% and 14% is reached at the expense of 3% and 4% active power consumption in the non-clocking and clocking 16-bit static-skew CLA adder, respectively.

**Keywords**: Low Power CMOS, Wireless Sensor Node, Energy-Efficient Circuits, Sub-Threshold Design, Forward Body Bias, Precharge-Evaluate Circuits, Low Leakage, Dual-Threshold Voltage, Static-Skew Circuits, Minimum Energy Consumption
# Contents

1 Introduction ............................................. 1
   1.1 Motivation ........................................... 2
   1.2 Embedded Microsystems ............................... 4
   1.3 Energy Aware Computation .......................... 6
      1.3.1 Hardware Software Co-design - System Level Approach .... 6
      1.3.2 Dynamic Energy Scaling - Circuit Level Approach ... 8
   1.4 Thesis Organization .................................. 11

2 Low-Leakage Circuit Techniques ....................... 15
   2.1 Introduction ......................................... 15
   2.2 Sources of Leakage Current ......................... 15
      2.2.1 Reverse-Biased Junction Leakage Current .......... 17
      2.2.2 Gate-Induced Drain Leakage Current .............. 18
      2.2.3 Gate Direct-Tunneling Leakage Current .......... 18
      2.2.4 Sub-Threshold Leakage Current .................. 18
   2.3 Effective Circuit Techniques to Control Leakage Power .... 21
      2.3.1 Dual-Threshold Voltage Technique ............... 21
      2.3.2 Sleep Transistor or Power Gating Technique .... 23
   2.4 Standby Leakage Power Performance Optimization - Circuit Level .. 26
      2.4.1 Static CMOS .................................... 27
      2.4.2 Single Rail Pass Transistor Logic ............... 27
      2.4.3 Differential-Cascode Voltage Switch with Pass-Gate Logic ... 29
      2.4.4 Simulation Results and Discussion ............... 30
   2.5 Standby Leakage Power Performance Optimization - Gate Level .... 32
      2.5.1 Simulations Results and Discussion ............... 33
Contents

4.4.1 Characteristics of PNFBB concept .......................... 103
4.4.2 Simulation Results and Discussion .......................... 106

5 Simulation Results ............................................ 113

5.1 Carry-Lookahead adder ....................................... 113
5.2 Design and implementation .................................. 116
  5.2.1 PG and Carry Blocks .................................... 116
  5.2.2 Simulation setup ........................................ 118
  5.2.3 Simulation Results and Discussion ....................... 121

6 Conclusion and Future Work .................................. 129
### Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{dd}$</td>
<td>Supply voltage</td>
</tr>
<tr>
<td>$V_{th}$</td>
<td>Transistor threshold-voltage</td>
</tr>
<tr>
<td>$I_{rev}$</td>
<td>reverse-biased junction leakage current</td>
</tr>
<tr>
<td>$I_{GIDL}$</td>
<td>Gate induced drain leakage current</td>
</tr>
<tr>
<td>$I_G$</td>
<td>Gate direct-tunneling leakage current</td>
</tr>
<tr>
<td>$I_{sub}$</td>
<td>Sub-threshold leakage current</td>
</tr>
<tr>
<td>$V_{bs}$</td>
<td>body to source voltage</td>
</tr>
<tr>
<td>$\phi_F$</td>
<td>fermi potential</td>
</tr>
<tr>
<td>$W_{dm}$</td>
<td>maximum depletion layer width</td>
</tr>
<tr>
<td>$\mu_0$</td>
<td>zero-bias mobility</td>
</tr>
<tr>
<td>$t_{ox}$</td>
<td>gate-oxide thickness</td>
</tr>
<tr>
<td>$C_{dm}$</td>
<td>capacitance of the depletion layer</td>
</tr>
<tr>
<td>$\epsilon_{ox}$</td>
<td>dielectric constant of $SiO_2$</td>
</tr>
<tr>
<td>$\epsilon_{si}$</td>
<td>permittivity of silicon</td>
</tr>
<tr>
<td>$N_A$</td>
<td>carrier concentration of intrinsic silicon</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>body-bias coefficient</td>
</tr>
<tr>
<td>$\epsilon_{ox}$</td>
<td>higher gate dielectric</td>
</tr>
<tr>
<td>$N_a$</td>
<td>lower channel implants or doping concentration</td>
</tr>
<tr>
<td>$\alpha'$</td>
<td>activity factor (the average number of transitions per clock cycle)</td>
</tr>
<tr>
<td>$C_{fs}$</td>
<td>fast surface state capacitance per unit area</td>
</tr>
<tr>
<td>$C_{ox}$</td>
<td>gate-oxide capacitance per unit area</td>
</tr>
<tr>
<td>$C_d$</td>
<td>Channel depletion region capacitance per unit area</td>
</tr>
<tr>
<td>$Q_c$</td>
<td>amount of charge stored on the capacitance</td>
</tr>
<tr>
<td>$n$</td>
<td>number of logic stages or logic depth</td>
</tr>
<tr>
<td>$C_s$</td>
<td>switching capacitance</td>
</tr>
<tr>
<td>$f$</td>
<td>delay chain frequency</td>
</tr>
<tr>
<td>$I_{Static}$</td>
<td>static-leakage current</td>
</tr>
<tr>
<td>$P_{AC}$</td>
<td>active-switching power</td>
</tr>
<tr>
<td>$P_{ST}$</td>
<td>static-leakage power</td>
</tr>
<tr>
<td>$V_T$</td>
<td>thermal voltage</td>
</tr>
<tr>
<td>$T$</td>
<td>temperature</td>
</tr>
<tr>
<td>$K$</td>
<td>Boltzmann’s constant</td>
</tr>
<tr>
<td>$q$</td>
<td>electronic charge</td>
</tr>
</tbody>
</table>
Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{eff}$</td>
<td>effective mobility</td>
</tr>
<tr>
<td>$W$</td>
<td>gate width</td>
</tr>
<tr>
<td>$m$</td>
<td>sub-threshold slope factor</td>
</tr>
<tr>
<td>$V_{dd_{min}}$</td>
<td>minimum operating supply voltage</td>
</tr>
<tr>
<td>$E_{min}$</td>
<td>minimum energy-per-cycle</td>
</tr>
<tr>
<td>$t_{d_{LVT}}$</td>
<td>intrinsic gate delays of low threshold gates</td>
</tr>
<tr>
<td>$t_{d_{HVT}}$</td>
<td>intrinsic gate delays of high threshold gates</td>
</tr>
</tbody>
</table>

List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMOS</td>
<td>Complementary Metal-Oxide Semiconductor</td>
</tr>
<tr>
<td>MEMS</td>
<td>Micro-Electro-Mechanical Systems</td>
</tr>
<tr>
<td>ITRS</td>
<td>International Technology Roadmap for Semiconductors</td>
</tr>
<tr>
<td>LSI</td>
<td>Large Scale Integration</td>
</tr>
<tr>
<td>DES</td>
<td>Dynamic Energy Scaling</td>
</tr>
<tr>
<td>SCMOS</td>
<td>Static CMOS</td>
</tr>
<tr>
<td>SRPL</td>
<td>Single Rail Pass Transistor Logic</td>
</tr>
<tr>
<td>SNOC</td>
<td>Sensor Node on Chip</td>
</tr>
<tr>
<td>AFBB</td>
<td>Asymmetric Forward Body Bias</td>
</tr>
<tr>
<td>LSTP</td>
<td>Low Standby Power</td>
</tr>
<tr>
<td>LOP</td>
<td>Low Operating Power</td>
</tr>
<tr>
<td>HLH</td>
<td>High-Leakage Header</td>
</tr>
<tr>
<td>LL</td>
<td>Low-Leakage</td>
</tr>
<tr>
<td>HL</td>
<td>High-Leakage</td>
</tr>
<tr>
<td>BTBT</td>
<td>Band-To-Band-Tunneling</td>
</tr>
<tr>
<td>DIBL</td>
<td>Drain-Induced Barrier Lowering</td>
</tr>
<tr>
<td>GIDL</td>
<td>Gate-Induced-Drain-Leakage</td>
</tr>
<tr>
<td>PUN</td>
<td>Pull-Up Network</td>
</tr>
<tr>
<td>PDN</td>
<td>Pull-Down Network</td>
</tr>
<tr>
<td>DTMOS</td>
<td>Dynamic-Threshold MOS</td>
</tr>
<tr>
<td>PDP</td>
<td>Power-Delay-Product</td>
</tr>
<tr>
<td>EDP</td>
<td>Energy-Delay-Product</td>
</tr>
<tr>
<td>MTMOS</td>
<td>Multi-Threshold CMOS</td>
</tr>
<tr>
<td>FO4</td>
<td>Fan-Out of 4</td>
</tr>
<tr>
<td>SR</td>
<td>Skew-Ratio</td>
</tr>
<tr>
<td>FBB</td>
<td>Forward Body Bias</td>
</tr>
<tr>
<td>RBB</td>
<td>Reverse Body Bias</td>
</tr>
<tr>
<td>ZBB</td>
<td>Zero Body Bias</td>
</tr>
<tr>
<td>PNFBB</td>
<td>Precharge Node based Forward Body Bias</td>
</tr>
<tr>
<td>CLA</td>
<td>Carry-look-Ahead</td>
</tr>
</tbody>
</table>
## List of Figures

1.1 The scaling trend of nominal transistor feature size and physical gate length dimensions [Cha05] ............................... 2
1.2 The increasing trend of transistors per microprocessor and on-chip memory requirements for computing applications [Cha05] .................................................. 3
1.3 General Sensor Node on a Chip Architecture ......................... 4
1.4 An adaptive Energy Harvesting Circuit [Ott02] ....................... 5
1.5 Power versus operating time of sensor node on a chip ................ 7
1.6 System-level hardware software co-synthesis flow .................. 7
1.7 The top-level architecture of Dynamic Energy Scaling technique ................................................................. 9
1.8 Temperature versus power-delay-product with forward body bias in ultra-low voltage regime ........................................ 10
2.1 Sub-threshold leakage power is dominant over the active power due to scaling down of physical gate length in the nanometer regime [INTEL] ..................................................... 16
2.2 Summary of n-channel MOSFET Leakage currents ............... 17
2.3 Drain-Induced Barrier Lowering - Decreasing Barrier height with the channel-length reduction and increase in drain-to-source voltage in an nMOS device [Nar06] ...................... 20
2.4 Dual-threshold voltage technique with low-threshold and high-threshold devices or gates are placed in the critical and non-critical path respectively. Here, the thick and thin-lines represent the low and high $V_{th}$ devices or gates respectively. ........................................ 22
List of Figures

2.5 Various configuration of sleep transistor or power gating techniques
   (a) Conventional power gating with low-threshold logic and high-threshold nMOS and pMOS SLEEP device
   (b) Only nMOS SLEEP device (c) Only pMOS SLEEP device and (d) Power gating with nMOS and pMOS body bias technique. 24

2.6 Dedicated LSI Blocks with pMOS Header ................................. 26

2.7 (a) Conventional static CMOS NAND with pMOS Header
   (b) & (c) HL-SRPL NAND & AND gates with output level restorer (d) HL-DCVSPG NAND/AND pMOS Header .......................... 28

2.8 Power-delay-product of HL, HLH and LL gates using SC-MOS, SRPL and DCVSPG logic circuits .............................. 30

2.9 Dual-threshold voltage technique in which low-threshold gates with pMOS Header and high-threshold gates are placed in the critical and non-critical paths respectively. .................. 32

2.10 Power-Delay-Product and standby-leakage power for SRPL and DCVSPG adder circuit .................................................. 33

3.1 Technology scaling according to ITRS roadmap predication ........... 36

3.2 (a) Theoretical minimum allowable supply voltage with respect to temperature and sub-threshold slope factor (b) Minimum switching energy-per-cycle exacerbates with temperature .................................................. 39

3.3 (a) Power consumption for a 13-stage adder decreases with supply voltage (b) The optimal minimum energy-per-cycle and the corresponding supply voltage (130nm technology) .......................... 42

3.4 High-speed MOSFET threshold voltage reduction and the corresponding bulk-current increase for different body bias voltage at elevated temperature (pMOS) ........................................ 45
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>Cross-section of Triple-well CMOS technology with active and static currents and the parasitic diodes are shown between the wells, body-to-source and body-to-drain nodes respectively</td>
</tr>
<tr>
<td>3.6</td>
<td>MOSFET threshold voltage and the corresponding drive-current for low leakage and high-speed device with wide-range of body bias voltage</td>
</tr>
<tr>
<td>3.7</td>
<td>Positive temperature dependence of Low-Leakage and High-speed devices in low $V_{dd}$ region at elevated temperature (between 25°C to 125°C)</td>
</tr>
<tr>
<td>3.8</td>
<td>The 13-Stage adder based delay chain with an enabling input block to trigger oscillation</td>
</tr>
<tr>
<td>3.9</td>
<td>Schematic cross-section of conventional static CMOS adder with minimum size devices</td>
</tr>
<tr>
<td>3.10</td>
<td>General architecture of body biasing in CMOS logic gates using external body bias circuitry</td>
</tr>
<tr>
<td>3.11</td>
<td>The delay chain frequency increases with FBB in the sub-threshold, near-threshold and above-threshold regions at elevated temperature (All data points are compared to $V_{bb}=0V$ to find out how many times the frequency increase w.r.t ZBB)</td>
</tr>
<tr>
<td>3.12</td>
<td>Performance improvement of delay chain with optimum FBB at different operating voltage, all data points normalized to conventional Zero Body Bias scheme</td>
</tr>
<tr>
<td>3.13</td>
<td>Performance improvement of 31-stage inverter based delay chain with optimum FBB and transistor sizing compared to conventional Zero Body Bias scheme</td>
</tr>
<tr>
<td>3.14</td>
<td>Simulated static noise margin $V_{OH}$ and $V_{OL}$ for an minimum size cascaded inverter with conventional zero body bias and forward body bias schemes</td>
</tr>
<tr>
<td>3.15</td>
<td>Relative static noise margin of FBB and Conv. ZBB at various supply voltages</td>
</tr>
</tbody>
</table>
List of Figures

3.16 (a) Conventional Zero Body Bias (b) Fixed Forward body bias with $V_{bb}= Optimal bias$ [External body bias generation] (c) Swapped Forward Body Bias with $V_{bb}=V_{dd}$ [Internal body bias generation] 61

3.17 Schematic cross-section of Forward Body Bias system to generate $V_{bsn}$ and $V_{bsp}$ respectively (Mas05) 62

3.18 (a) Asymmetric FBB with well-contact tied together (b) Asymmetric n-well Forward Bias - only pull-up devices are forward biased ($V_{AFB}=0V$) (c) Asymmetric p-well Forward Bias - only pull-down devices are forward biased ($V_{AFB}=V_{dd}$) 63

3.19 The 13-stage adder delay chain frequency improvement and the corresponding total power consumption with symmetric FBB and conventional zero body bias schemes at the elevated temperature 65

3.20 The 13-stage adder delay chain frequency is improved with asymmetric FBB compared to the conventional ZBB scheme at the elevated temperature 67

3.21 The change in power consumption with asymmetric FBB compared to conventional ZBB scheme at the elevated temperature 68

3.22 Minimum energy-per-cycle for conventional and swapped FBB schemes at elevated temperature (Symmetric FBB) 70

3.23 Minimum energy-per-cycle for n-well, p-well forward bias and conventional ZBB schemes (Asymmetric FBB) 71

3.24 Energy-per-cycle and the corresponding performance improvement of 13-stage adder based delay chain for symmetric and asymmetric FBB schemes at the elevated temperature 72

3.25 Energy-Delay-Product and the respective frequency of 13-stage ADDER delay chain using symmetric and asymmetric FBB schemes at the elevated temperature 74

3.26 Worst-case analysis of power and frequency with different process corners using 13-stage adder based delay chain 76
3.27 Worst-case analysis of 13-stage adder based delay chain in terms of energy-per-cycle, supply voltage and frequency of the delay chain ........................................ 77

3.28 $(3\sigma/\mu)$ 31-stage inverter based delay chain showing the delay and active energy variation with optimal FBB and various transistor dimensions ........................................ 80

3.29 $(3\sigma/\mu)$ 13-stage adder delay chain showing the delay and active energy variation at sub-threshold and above-threshold region for various body bias configuration ........... 82

4.1 (a) Conventional zero body bias (b) and (c) Symmetric configuration with $V_{bs}$ = fixed optimal value and $V_{bs}$ = $V_{dd}$ respectively .................................................. 87

4.2 (a) Asymmetric configuration with single bias voltage (b) and (c) n-well and p-well forward bias with well-contacts connected to $V_{ss}$ and $V_{dd}$ rails respectively ............ 88

4.3 Fan-out (FO3) of 3 Inverter based 51-stage delay chain ....... 89

4.4 The change in frequency of 51-stage delay chain with asymmetric bias voltage compared to conventional zero body bias scheme (all data points are normalized to conventional ZBB) 89

4.5 The change in active and static power (average) of 51-stage delay chain with asymmetric bias voltage compared to conventional zero body bias scheme .......................... 91

4.6 The change in active and static power (average) of 51-stage delay chain with asymmetric bias voltage compared to conventional zero body bias scheme .......................... 93

4.7 Circuit topology and its connection (a) Clock-less Skewed NAND gate (Slow precharge) (b) Clock-less Skewed NOR gate (c) Clocked Skewed NAND gate (Fast precharge) [Skew Ratio = 3] .................................................. 95

4.8 Pulsed static CMOS - precharging the output node by input pattern without any clock scheme ........................................ 96
List of Figures

4.9 A chain of skewed gates between two latches A and B. Alternate low-skew and high-skew gates are indicated by the arrows ........................................ 97

4.10 Propagating an inverter through logic gate (a) without fanout (b) with fanout .................................................. 98

4.11 Forward propagation of intermediate inverters towards primary outputs .................................................. 98

4.12 Backward propagation of inverters to obtain inverter-free logic (a) logic with NAND/NOR/AND/OR/NOT gates (b) logic with AND/OR/NOT gates (c) propagating inverters back towards primary inputs without any logic duplication (d) combining the inverters trapped at intermediate fanouts (e) inverter-free logic after propagating the inverters back towards primary inputs with logic duplication .................. 100

4.13 Precharge-Node based Forward Body Bias (PNFBB) technique ................................................................. 101

4.14 High-speed static skewed inverter with PNFBB technique ................................................................. 102

4.15 Low-leakage static skewed inverter with PNFBB technique ................................................................. 103

4.16 Fan-out of 4 inverter delay versus Skew ratio ................................................................. 104

4.17 Comparison of low skew 3-input NAND evaluation delay with conventional zero body bias and the proposed method for different output load capacitance and skew ratio ........ 105

4.18 Dual-threshold voltage technique with low and high threshold voltage devices placed in the critical and non-critical paths respectively ................................................................. 106

4.19 (a) & (b) PNFBB 2-input skewed AND gate with static inverter (c) PNFBB 2-input Domino AND gate with static inverter (Skew Ratio =3) and (d) Simulation setup with equal sized buffers at both sides of complex AOI logic .......... 107

4.20 (a) Supply voltage versus Energy per cycle and (b) Supply voltage versus Evaluation Delay for PNFBB and Dual threshold voltage technique ................................................................. 108

xiv
4.21 (a) Supply voltage versus active switching power and (b) supply voltage versus static power for PNFBB and Dual threshold voltage technique .............................................. 109

4.22 The active energy-per-cycle and evaluation delay of PNFBB fast precharge static skewed gate for various supply voltages and Skew ratio ................................................................. 110

5.1 Generic block diagram of Carry-Look Ahead structure (a) Conventional CLA structure (b) Modified CLA structure ................................................................. 115

5.2 Gate-level implementation of modified static skewed 4bit group PG and carry blocks using dynamic Manchester carry chain ......................................................... 116

5.3 Transistor level circuit implementation of modified static skew 4bit group PG and carry blocks using dynamic Manchester carry chain (optional clock shown by dotted lines). Note that low-skew NAND gates and high-skew inverters are connected in PNFBB approach ................................................................. 117

5.4 Modified circuit implementation of static skew 4bit group PG and carry blocks using static manchester carry chain. The low-skew and high skew gates are connected in PNFBB technique ................................................................. 119

5.5 Experimental set up of 16-bit static skew adder with modified CLA with input buffers and output load stages ................................................................. 120

5.6 Schematic cross-section of 16-bit static skew adder with modified CLA structure. The optional clock to precharge the output node is shown in solid dotted lines ................................................................. 121

5.7 Simulated waveform of sub-block 4-b CLA of 16-b skew CLA adder using proposed PNFBB technique ................................................................. 122

5.8 Static skew 16-b CLA adder using dynamic manchester carry chain using clocking and non-clocking scheme for conventional zero body bias and the proposed technique ($V_{dd} = 0.5V$) 123
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.9</td>
<td>Evaluation delay and the corresponding active energy reduction for 16-b static skew adder using proposed PNFBB technique</td>
<td>124</td>
</tr>
<tr>
<td>5.10</td>
<td>Static skew 16-b CLA adder with static Manchester carry chain using conventional zero body bias and the proposed technique (non-clocking version)</td>
<td>126</td>
</tr>
</tbody>
</table>
# List of Tables

1.1 Characterization of simulated circuit for different supply voltage and forward body bias (25°C) ........................................ 11

2.1 Standby-leakage power of SCMOS, SRPL & DCVSPG logic circuits ................................................................. 31

2.2 Power-Delay-Product and standby-leakage power for different schemes using SCMOS full-adder circuit ................. 34

3.1 Mode of operation of parasitic diodes with applied body bias in the triple-well CMOS process technology. Forward bias means positive voltage, Reverse bias means negative voltage and zero bias means zero voltage across the nodes ........ 48

3.2 Reduction in Threshold voltage and the respective drain currents when $V_{GS} = V_{dd} (I_{on})$ and $V_{GS} = 0V (I_{off})$ with 0.5V FBB 48

3.3 Mean ($\mu$), standard deviation ($\sigma$), delay and power variability ($\frac{\sigma}{\mu}$) for conventional zero body bias, asymmetric FBB and symmetric FBB schemes respectively in the above-threshold voltage region (0.5V $V_{dd}$) ........................................ 81

3.4 Mean ($\mu$), standard deviation ($\sigma$), delay and power variability ($\frac{\sigma}{\mu}$) for conventional zero body bias, asymmetric FBB and symmetric FBB schemes respectively in the sub-threshold voltage region (0.2V $V_{dd}$) ........................................ 81

4.1 Power and performance comparison of 2-input NAND and NOR with conventional Zero body bias, n-well Forward bias and p-well Forward bias respectively ($V_{dd} = 0.5V$) ............. 92
List of Tables

4.2 Comparison of fast precharge complex static skew AOI evaluation delay and energy-per-cycle with zero body bias scheme (skew ratio=3) .......................................................... 111

5.1 UMC 130nm technology high speed device parameters (W/L=10/0.12) 119
1 Introduction

Minimization of energy consumption in the battery powered portable applications is always a challenging task in the scaled down CMOS technologies. The increasing demand for low power and high performance VLSI systems acts as major driving force for technology scaling and higher energy efficiency. The tremendous success of the semiconductor industry over the last years causes significant change in our daily lifestyle. Electronic circuits are widely used almost in each and every part of our life starting from computers to automobile systems. Integrated circuits are still advancing at a rapid pace, in a remarkably enduring extension of the trend known as "Moore’s Law", based on a projection by Gordon Moore [Moo65]. Ultra modern embedded devices have continuously overcome postulated limits to technological progress. Therefore, transistors are now being manufactured with gate dimensions well below 100nm [Cha05] as shown in Fig. (1.1).

Tremendous shift in the integrated circuits occurred in 1970’s and 1980’s, like the transitions from bipolar to pMOS, nMOS and CMOS technologies. In recent years, the number of transistors on microprocessor chips has grown at a faster rate with increasing use of on-chip cache memory which is rising above billion transistor level as shown in Fig. (1.2). This causes significant increase in performance and power efficiency with smaller device geometry. Ideally, CMOS has no static power consumption and therefore the total power consumption is only due to the active switching activity of the circuit nodes. Voltage scaling is the most effective method to reduce active power consumption, in which the operating voltage is dropped from 5V to 1V and this reduces the power consumption significantly (≈25X). On the other hand, the threshold voltage has to be lowered to improve performance but it increases the sub-threshold leakage current exponentially. Also, the gate-leakage current increases because of smaller gate-oxide thickness, but which is highly
1 Introduction

Silicon Technology Reaches Nanoscale

Figure 1.1: The scaling trend of nominal transistor feature size and physical gate length dimensions [Cha05]

beneficial for improving performance. Due to the substantial increase in the leakage current, the static power is expected to exceed the active power consumption unless effective measures are taken to minimize the leakage power. Hence, industry and academic institutions are working to innovate novel process, device and circuit design techniques for minimizing the total power consumption.

1.1 Motivation

Minimizing power consumption is always a primary requirement for battery driven embedded low-power processor design. The increasing prominence of portable devices has become a fundamental driving factor in the design of new computational CMOS elements. As the target shifts from desktop computing to mobile applications, rethinking of design optimizations is required with ever-increasing performance goals and for extending the usability of portable gadgets. The power consumption arises as a third axis in the optimization space among other design metrics because it reduces the package size, extends battery life time and minimizes the overhead of thermal management which result in low costs and higher reliability. Thus, a power efficient design means the system minimizes peak demands and improves its
1.1 Motivation

Moore’s Law

Figure 1.2: The increasing trend of transistors per microprocessor and on-chip memory requirements for computing applications [Cha05]

operating frequency [Moy01]. However, the performance improvement with limited power resources is a challenging task in the low power design. Supply voltage scaling is the most effective way of reducing power consumption and it has been practically demonstrated in a number of designs. The supply voltage reduction with fixed threshold voltage provides quadratic reduction in dynamic power consumption at the cost of reduced performance. This performance penalty is tolerable in many applications like wireless sensor nodes in which minimum power consumption is the primary concern. The main driving force for ultra-low voltage low-power design is the ubiquitous computing applications like wireless sensor nodes in which the distributed nature of sensor network results in huge energy constraints on the sensor nodes. Even high-performance applications benefit from voltage scaling, since the supply voltage could be reduced to lowest operating voltage during idle time in which the processor performs simple (or) no operations. Regardless of the application, the energy reduction by scaling supply voltage improves lifetime of the battery source whenever the performance demands are not of primary concern. Hence, the energy efficient design improves the lifetime of its power source, while it attempts to meet the throughout and peak performance requirements of the overall applications.
1 Introduction

Figure 1.3: General Sensor Node on a Chip Architecture

1.2 Embedded Microsystems

A wireless sensor network is a cluster of sensor nodes communicating with each other for collecting, processing and distributing data [Aky02]. Each sensor module consists of sensors, data converters, a processor, power management unit, radio for wireless communication and networking. Wireless sensor networks have a wide range of applications like video surveillance, industrial and manufacturing automation, distributed robotics, smart textiles, building, structure as well as environmental monitoring etc.

The general architecture of a wireless sensor node is shown in Fig. (1.3) and it consists of sensors, application specific dedicated hardware to process and control the data, an RF modem with DSP engine to send and receive messages between nodes or base station and the power management block. Although, most of the micro-controllers have various power-down and idle modes they still tend to consume large amount of power. Even at a moderate speed of 1 MHz, it require microwatts of
1.2 Embedded Microsystems

Figure 1.4: An adaptive Energy Harvesting Circuit [Ott02]

power in the power down and milliwatts in the active mode. However, a simple $1cm^3$ Lithium battery can provide $10\mu W$ of power for five years without any interruption [Cal05]. To extend the life time, on-board self renewable energy source is necessary to continuously replenish the energy consumed by the autonomous embedded microsystems. State-of-the-art micro-electromechanical-system (MEMS) generators and transducers can be used to harvest the energy from the ambient source (light, thermal gradients, motion, stress, vibrations and etc) and converting into electrical energy. Since most of the power sources will deliver lower levels of useable energy, systems that find ways to maximize the effectiveness would be interesting. Geffrey K. Ottman et.al proposed an adaptive technique to harvest electrical energy from a mechanically excited piezoelectric element which is shown in Fig. (1.4). It reveals that use of the adaptive dc-dc converter increases power transfer by over 400% as compared to when the dc-dc converter is not used. Another attractive power source might be feasible in some applications using RF energy transmission [Man05] which depends on the size of the device, the available area for an antenna and the distance that has to be covered, by adapting one of these method quite reasonable amount of power can be transmitted. This power source has been effectively demonstrated in many medical implants as well as in RF-IDs. Similarly, there are several other methods which have been proposed to prolong the life span of autonomous microsystems from the renewable energy harvesting mechanism.

Hence recharging or storing the energy from the ambient source has to be utilized in the effective way to improve the efficiency of the sensor node. To improve this energy efficiency, the more careful power management techniques has to be adapted
in the computation and communication units during active and standby mode of operation.

1.3 Energy Aware Computation

Power optimization can be done at various levels of abstraction viz., behavioral, architectural and logical or physical levels. The behavioral and architectural level optimization can bring 10% to 90% power savings [Gai98] [Hak00] [Hak99]. Dedicated hardware for low level tasks and careful use of software for high level tasks reduces the energy at the architectural level while adaptive supply and threshold voltage schemes reduce the power at the logic and circuit level. To prolong the lifetime of the sensor nodes, all aspects of the node should be energy efficient. This energy requirement depends on the state of operation, like standby or active mode and the nature of applications. However, in many applications the standby mode will be longer than the active operating time which is depicted in Fig. (1.5) and this requirement causes the sensor node to have different operating modes. In many application from computing to mobile devices, the wake-up/active time is much smaller than the sleep/power-down time. This forces these systems to be also optimized for a higher standby power consumption. A number of different sleep modes only maintain the blocks active which are required for a specific background task. Dedicated sensor data acquisition hardware collect data or compare them to critical values without the intervention of the micro-controller. Therefore, the new methodologies and concepts are mandatory due to growing complexities of embedded wireless sensor node. At first, the design process must be shifted towards the higher levels of abstraction and hence the focus is set upon a system-level design methodology.

1.3.1 Hardware Software Co-design - System Level Approach

In order to efficiently utilize the limited energy resources available in a sensor node, one has to optimize its key design parameters to improve its efficiency. This can be only possible by making decisions on system-level design about its hardware and software (operating system and applications) architecture. To achieve this objective, the co-synthesis needs to address four fundamental design problems such as architec-
Figure 1.5: Power versus operating time of sensor node on a chip

Figure 1.6: System-level hardware software co-synthesis flow
1 Introduction

ture allocation, application mapping, activity scheduling, and energy management. Figure (1.6) shows the co-synthesis flow in diagrammatic form [Ele04]. The energy management techniques that we can apply to wireless sensor nodes is dynamic voltage scaling and transistor body biasing. To avoid unnecessary high switching activity, the application should be carefully partitioned into cores so that, they can be selectively switched off. This technique is particularly suitable for designs in which no clock gating is applied. [Dav97] developed one approach that targets the reduction of power dissipation throughout the co-synthesis process. They developed a constructive algorithm based on energy levels, which makes the mapping of tasks energy sensitive. [Luo00] developed a combined scheduling technique for periodic executing tasks with dependencies as well as aperiodic tasks. Dynamic voltage scaling is considered in this scheduling context and dependent tasks are scheduled statically. The scheduling of aperiodic tasks as well as utilization of dynamic voltage scaling is improved by distributing available slack time among the processing elements of a distributed architecture. Using the methods proposed above, one can optimize each step for energy, area and cost, then evaluate the design and if necessary go back to different steps in the flow until the requirements are satisfied.

1.3.2 Dynamic Energy Scaling - Circuit Level Approach

In the ubiquitous computing era, the standby leakage power starts dominating over the active switching power due to scaled down technologies [Bor02]. To design an energy efficient LSI, minimizing both the active switching power and standby leakage power without much performance penalty is a challenging task. In case of autonomous sensor nodes, the energy reduction is usually a major concern over performance improvement. Therefore, the energy efficient LSI is required to maximize the useful lifetime of the battery source, which is accomplished by adjusting both the supply and threshold voltage depending on the mode of operation and the system workload [Kur96]. Hence, we made an preliminary study in adapting the supply voltage and the threshold voltage by constantly maintaining the performance of the system for burst mode computation. The change in the threshold voltage and the corresponding change in the supply voltage results in a constant performance with lower active and standby leakage power. Since the threshold voltage has very strong
1.3 Energy Aware Computation

dependency on the ambient temperature and process parameter fluctuation, it could be altered to maintain fixed performance targets by scaling down the supply voltage. By this approach, both the active and standby leakage power could be reduced to least possible value at an optimal performance (fixed), with the aid of body bias voltage.

![Temperature Sensor](image1.png)

![Performance Tracking System](image2.png)

![Body Bias Generator](image3.png)

![Voltage Regulator](image4.png)

![Feedback Loop](image5.png)

**Figure 1.7: The top-level architecture of Dynamic Energy Scaling technique**

The conventional dynamic voltage scaling method [Bur00] reduces the supply voltage based on the performance demands, but it have not taken into account the standby leakage power and process parameter variations. Hence, we made an initial attempt in which the energy requirements are altered dynamically based on the fixed performance targets and process parameter variations as shown in Fig. (1.7). During active mode of operation, the positive body bias is applied to the high-threshold voltage LSI which maintains the performance at fixed rate with reduced supply voltage. During standby mode of operation, the applied body bias is withdrawn to bring the LSI to initial state (high-threshold voltage state) which maintains the leakage current under control. By this technique, the lower active power is achieved while maintaining the same performance level as that of higher supply voltage, with lower leakage power.

Several different hardware with various supply voltages operated at elevated temperature are considered and the circuit simulation has been done using 180nm mixed-mode process technology and the simulated values are as shown in Fig. (1.8). Since, the threshold voltage and mobility have strong temperature dependence which affects the power, delay and the power-delay-product. As evident from Fig. (1.8), the
1 Introduction

temperature dependence w.r.t power-delay-product is larger at higher supply voltage (1V), however it is relatively lower at reduced supply voltage due to positive temperature coefficient (smaller delay). If the comparison is made between 0.8V and 0.6V with zero body bias (ZBB) an significant power reduction of 13.7 times is obtained at the expense of severe delay penalty (3.8 times delay increase). To minimize this, the forward body bias (FBB) is applied which reduces the delay penalty with extra power due to junction leakage current. In order to find the optimum between these boundaries, an optimal FBB is applied at 0.6V supply voltage, it reduces both the power-delay-product and delay penalty compared to 0.8V with ZBB, respectively. It shows that the FBB allows a variation in the performance level without having to

Figure 1.8: Temperature versus power-delay-product with forward body bias in ultra-low voltage regime

continue the supply voltage, which could be associated with a much higher overhead in the voltage regulator and the level-shifters at the boundaries of different supply voltages. The overhead for FBB is minimal since the bias voltage can be extracted from the supply voltage with a simple resistor ladder. During the standby mode, withdrawing the body bias voltage maintains high threshold voltage state, which in turn reduces the standby leakage power of the sensor node. In addition, the lower supply voltage reduces the DIBL effect which will increase the threshold voltage and hence both the active and standby leakage power has been reduced considerably.
Table 1.1: Characterization of simulated circuit for different supply voltage and forward body bias (25°C)

<table>
<thead>
<tr>
<th>DesignMetric</th>
<th>0.8V_{dd} (Zero Body Bias)</th>
<th>0.6V_{dd} (Forward Body Bias)</th>
<th>0.6V_{dd} (Zero Body Bias)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power (nW)</td>
<td>811.5</td>
<td>280.8</td>
<td>58.83</td>
</tr>
<tr>
<td>Delay (pS)</td>
<td>131.9</td>
<td>276.2</td>
<td>496.2</td>
</tr>
<tr>
<td>Power-delay-product (fJ)</td>
<td>0.107</td>
<td>0.07</td>
<td>0.02</td>
</tr>
</tbody>
</table>

So far, we have analyzed the system level hardware-software co-design approach and dynamic energy scaling methodology to design energy efficient computing blocks. Furthermore, we analyzed the challenges faced by designers and developers to deliver highly energy efficient wireless sensor node. We discussed many strategies of electronic power reduction for sensor nodes at various levels of abstraction. It has been shown that application specific hardware for the low-level tasks and software for high-level tasks together with dynamic energy scaling reduces the power consumption at the architectural and circuit level respectively. Adapting these techniques to a wireless sensor node to operate in ultra-low power mode, makes energy harvesting a possible choice.

1.4 Thesis Organization

Energy efficient design for embedded microsystems requires significant design optimization at all levels of design hierarchy. In this work, the ultra-low voltage circuit techniques and the methods to obtain highly energy efficient design by reasonably meeting the performance targets are presented. At first, in this section we discussed the methods to optimize power at different levels of design hierarchy i.e., from architecture to circuit level targeting the power optimized wireless sensor node architecture. The primary focus of this work is to develop energy reduction strategies which is described in detail in the following paragraphs.

Low power and low leakage circuit techniques are addressed in Chapter 2. At first, it evaluates different leakage current sources in the nanometer regime and
then presents effective methods for minimizing the dominant sub-threshold leakage currents. Since, the sub-threshold leakage is dominant over other sources, it has primary focus throughout our discussion. Further, it address the strategies to determine the best optimal choice by combining the state-of the art logic circuit families with effective leakage reduction techniques. Finally, the effective standby and active leakage reduction techniques like power gating or sleep transistor and dual-threshold voltage respectively are addressed, at different levels of abstraction.

Energy efficient design is extensively discussed in Chapter 3. In this, the low voltage and low energy design methodologies are widely addressed in the ultra-low supply voltage ($V_{dd} \leq 0.5V$) with body bias approach. Forward body bias is introduced in the ultra-deep supply voltage region to obtain optimal trade-offs between the energy-per-cycle and the performance design metric. Further, the various categories of body bias configurations like symmetric and asymmetric schemes are discussed in detail from the perspective of both power and area design metric. The characteristics of symmetric and asymmetric forward body bias schemes in the ultra-low voltage regime are addressed in terms of noise margin, device width ratio and etc. Finally, the minimum energy-per-cycle is analyzed in the presence of process parameter fluctuation with which we discuss several approaches to mitigate the process variations with the help of body bias and device ratio in the ultra-low voltage region.

Ultra-low voltage circuit design with novel fine-grained self-adaptive body bias approach is presented in Chapter 4. It discusses the asymmetric forward body bias scheme for optimum energy-performance point and proposed the precharge-node based forward body bias technique for domino-like precharge-evaluate logic families. Further, the static and dynamic precharge-evaluate logic circuits are evaluated for conventional body bias approaches for comparison. Dual-threshold voltage techniques are addressed in the ultra-low voltage regime, and the results are compared with the proposed PNFBB approach. The synthesis of precharge-evaluate logic circuits and the simulation results are discussed with existing methods using an industrial 130nm process technology.
Chapter 5 presents the final results of the proposed method with a 16-bit adder using modified Manchester static and dynamic CLA structure. To determine the energy efficient adder, an extensive simulation has been performed with symmetric and asymmetric body bias configuration. Finally, chapter 6 summarizes the conclusion and proposes ways to further extend this research work.
1 Introduction
2 Low-Leakage Circuit Techniques

2.1 Introduction

In the nanometer regime, the static leakage power is becoming a major portion of the total power consumption in the battery operated embedded portable applications. Many recent research results that 40% or even higher percentage of the total power consumption is due to the transistor leakage current [Cha92] [Kim03]. This increased leakage power not only influences the total power consumption but it also minimizes the design margin due to strong dependence on process parameter fluctuations. This trend will continue to a large extent unless the effective methods are taken to bring leakage under control. It shows that the designers face critical challenges in choosing the appropriate logic circuit topologies with the state-of-the-art leakage mitigation techniques in order to optimally balance between the power and performance trade-offs. This section primarily focus on various sources of leakage current and methods of minimizing the active and standby leakage power in conventional CMOS digital circuits. In addition, it also evaluates different static logic circuits with effective active and standby leakage power reduction techniques for highly energy efficient design.

2.2 Sources of Leakage Current

Figure (2.1) illustrates the dominance of leakage power compared to the active switching power, when the physical gate length continuously scales down to improve energy efficiency of CMOS devices. With this rapid progress in semiconductor technology, the chip density and the operating frequency has increased, making the power consumption in battery-operated portable devices a major concern [Bor02]
2 Low-Leakage Circuit Techniques

Figure 2.1: **Sub-threshold leakage power** is dominant over the active power due to scaling down of physical gate length in the nanometer regime [INTEL]

[Ped05] [Viv01]. The primary objective of low-power design for battery-powered devices is to extend the battery life time while meeting the performance demands. Thus, the supply voltage scaling is perhaps the most preferable and effective way of saving power due to the quadratic dependency of digital active power on the supply voltage. However, it reduces the performance since the circuit driving capability is also reduced. Thus, the lowering of supply voltage helps in reducing the active power consumption of CMOS logic gates at the expense of performance loss. To overcome this negative effect, the threshold voltage has to be scaled down to maintain the performance constant. This causes the transistor leakage current to increase, which further increases the static power consumption. It shows that, there is a trade-off between the static leakage power and active power for a given application, leading to methods of selecting the threshold voltage and supply voltage for reaching fixed performance targets [Jay05]. One possible way to mitigate this task for the low energy LSI systems, is to exploit adaptive supply and threshold voltage methods to meet both the power and performance budget. Though the supply voltage and device dimensions scale down to minimize the active power, the static leakage power becomes a major bottleneck compared to the active power due to lowering of threshold voltage and increasing chip size. Thus, primarily we analyze the sources of leakage currents and then evaluate the possible approaches to minimize the sub-threshold
2.2 Sources of Leakage Current

leakage current in the nanometer era.

Figure (2.2) shows the major sources of leakage currents in the n-channel MOSFET [Nar06] [Roy03], which is categorized as 1. Reverse-biased junction leakage current \(I_{rev}\) 2. Gate induced drain leakage current \(I_{GIDL}\) 3. Gate direct-tunneling leakage current \(I_G\) and 4. Sub-threshold leakage current \(I_{sub}\). In the following sections, we will briefly discuss each of these components.

2.2.1 Reverse-Biased Junction Leakage Current

The reverse-biased junction leakage current \(I_{rev}\) occurs either from the source or drain to the substrate region through the reverse biased PN junction diodes when the device is turned OFF. Its mainly due to either minority carrier diffusion/drift near the depletion region or electron-hole pair generation in the depletion region of the junction [Ped05]. When both n-region and p-region are heavily doped, as in the case of some advanced MOSFETs, there is also junction leakage due to band-to-band tunneling (BTBT). For example, in the case of simple inverter with a low input, the drain-to-source voltage of the nMOS is the same as that of the supply voltage. This results in a reverse biased junction leakage current which flows from the drain to the substrate of the nMOS device. The magnitude of this leakage current is a function of diffusion area, temperature and the leakage current density, which in turn depends
on the doping concentration of both N and P-type regions. Its contribution to the total leakage is negligible compared to the other leakage currents.

### 2.2.2 Gate-Induced Drain Leakage Current

Gate induced drain leakage current ($I_{\text{GIDL}}$) arises in the high electric field under the gate/drain overlap region causing deep depletion, and effectively thinning out the depletion width of the drain to the well PN junction. Carriers are generated into the substrate and drain from the direct band-to-band tunneling, trap assisted tunneling, or a combination of thermal emission and tunneling. The thinner oxide thickness $T_{\text{ox}}$ and higher $V_{dd}$ causes a higher potential between the gate and drain which enhances the electric field dependent GIDL.

### 2.2.3 Gate Direct-Tunneling Leakage Current

Gate direct-tunneling leakage current ($I_g$) flows from the gate through the gate oxide material into the substrate. For oxide layers thicker than 3 to 4 nm, this current results from the Fowler-Nordheim tunneling of electrons into the conduction band of the oxide layer under a high applied electric field across the oxide layer. This tunneling typically lies at a higher field strength and has a constant slope for $E_{\text{ox}} > 6.5\text{MV/cm}$. This is presently not an issue but could surpass subthreshold and DIBL as a dominant leakage mechanism in the future as oxides get thinner. An effective approach to overcome the gate leakage currents while maintaining excellent gate control is to replace the currently used silicon dioxide gate insulator with high-K dielectric material such as $\text{TiO}_2$ and $\text{Ta}_2\text{O}_5$. Use of the high-k dielectric will allow a less aggressive gate dielectric thickness reduction while maintaining the required gate overdrive at low supply voltages.

### 2.2.4 Sub-Threshold Leakage Current

As supply voltage scales down with process technology, the threshold voltage has to be reduced to sustain 30% gate delay improvement. This threshold voltage reduction causes a significant increase in the sub-threshold leakage current ($I_{\text{sub}}$). It is the drain-to-source current of a transistor operating in the weak inversion region. Unlike the strong inversion region in which the drift current dominates, the subthreshold
2.2 Sources of Leakage Current

Conduction is due to the diffusion current of the minority carries in the MOSFET devices. In case of simple inverter with low input, there is a substantial amount of sub-threshold leakage current flowing in the off-state nMOS device due to higher drain-to-source voltage. This portion will vary depending on several factors such as the threshold voltage, temperature, supply voltage and transistor width. Among this, the threshold voltage reduction plays a major role in the modern device off-state leakage currents. This sub-threshold current can be expressed as follows [Tau98]

\[ I_{ds} = \mu_0 C_{ox} \frac{W}{L} (m - 1)(V_T)^2 e^{(V_s-V_{th})/mV_T}(1 - e^{-V_{DS}/V_T}) \]  

(2.1)

where

\[ m = 1 + \frac{C_{dm}}{C_{ox}} = 1 + \frac{\epsilon_{si}}{W_{dm}} = 1 + \frac{3t_{ox}}{W_{dm}} \]  

(2.2)

where \( V_{th} \) is the threshold voltage, and \( V_T \) is the thermal voltage. \( C_{ox} \) is the gate oxide capacitance; \( \mu_0 \) is the zero bias mobility; and \( m \) is the subthreshold swing coefficient. \( W_{dm} \) is the maximum depletion layer width, and \( t_{ox} \) is the gate oxide thickness. \( C_{dm} \) is the capacitance of the depletion layer.

As the device size and supply voltage are scaled down to enhance performance, power efficiency and reliability, the subthreshold characteristics may limit the scalability of the supply voltage. Hence, it is highly desirable to have a subthreshold swing as small as possible since this is the parameter that determines the amount of voltage swing necessary to switch a MOSFET from OFF to ON state (typical values for bulk CMOS devices are 70-110 mV gate-source voltage for a decade in current difference; the theoretical lower bound is 60 mV/decade). This is especially important for modern MOSFETs with supply voltages reaching sub-1V region. To minimize this parameter, the thinnest possible gate oxide (since it increases oxide capacitance) and the lowest possible doping concentration in the channel (since it decreases depletion capacitance) must be used. Moreover, the higher temperature results in larger subthreshold swing value, and hence, an increase in the subthreshold leakage current.

In long channel devices, the influence of source and drain on the channel depletion layer is negligible. However, as channel lengths are reduced, overlapping source and drain depletion regions cause the depletion region under the inversion layer to
Figure 2.3: Drain-Induced Barrier Lowering - Decreasing Barrier height with the channel-length reduction and increase in drain-to-source voltage in an nMOS device [Nar06]

increase. The wider depletion region is accompanied by a larger surface potential, which attracts more electrons to the channel. Therefore, a smaller amount of charge on the gate is required to reach the onset of strong inversion with smaller threshold voltage. This effect is worsened with a larger bias on the drain, since the depletion region becomes wider. To be precise, when a high drain voltage is applied to a short-channel device, it lowers the barrier for electrons between the source and channel, resulting in further decrease of threshold voltage. The source then injects carriers into the channel surface (independent of gate voltage), causing an increase in $I_{OFF}$. This phenomenon, which can be thought of as a lowering of $V_{th}$ as $V_{DS}$ increases, is the DIBL effect as shown in Fig. (2.3). There is yet another phenomenon known as the $V_{th}$-rolloff whereby the $V_{th}$ of a MOSFET decreases as the channel length is reduced. In such a case, the subthreshold swing parameter degrades and the impact of drain bias on $V_{th}$ increases. Finally, there is the well-known body effect, which causes an increase in $V_{th}$ as the body of the transistor gets reverse-biased. Thus, combining above mentioned leakage currents altogether, we can write the total leakage current as,

$$I_{LEAK} = I_{sub} + I_{rev} + I_{GIDL} + I_g$$

(2.3)

Among this, reverse-biased junction leakage and GIDL current is maximized when the drain-to-substrate bias is equivalent to the operating voltage. When the device
is switched off \( (V_{gs} = 0V) \), the off-state leakage current comprises only \( I_{sub}, I_{rev} \) and \( I_{GIDL} \). The gate leakage occurs only when the gate potential is at a high voltage with respect to the source and substrate.

### 2.3 Effective Circuit Techniques to Control Leakage Power

One major challenge with technology scaling is the rapid increase in sub-threshold leakage power, hence it becomes crucial to identify effective techniques to reduce this leakage component. In many portable applications, there are mostly two types of operation such as active mode in which the system performs useful tasks and standby or idle mode, respectively [Jay05a]. The leakage power consumed during active mode is considered as active leakage power and during standby mode is called standby leakage power.

There are several techniques which are widely used to reduce both active and standby leakage power which will be discussed in the following sections. We primarily focus on the dual-threshold voltage technique [Kao00] and the power-gating or sleep transistor technique [Mut95] to effectively mitigate the leakage during active and standby mode of operation.

#### 2.3.1 Dual-Threshold Voltage Technique

A promising technique for reducing leakage power during both active and standby mode while attaining the high performance during active mode, is to utilize multi-threshold devices in the circuit implementation. In this case, the high-threshold voltage and the low-threshold voltage devices or logic gates are placed in the performance non-critical and critical paths respectively, to suppress the active leakage component with minimal performance penalty. To achieve this, there are several methods proposed (Fig. 2.4) to mix low and high \( V_{th} \) devices or gates for the circuit and gate level implementation, to find the optimum between performance and power savings.

However, there are many design issues and challenges which have to be overcome from process fabrications steps to the state-of-the-art CAD tools. For instance, de-
Figure 2.4: Dual-threshold voltage technique with low-threshold and high-threshold devices or gates are placed in the critical and non-critical path respectively. Here, the thick and thin-lines represent the low and high $V_{th}$ devices or gates respectively.
vice level multi-threshold placement imposes several challenges in both design process and the complexity of using available EDA tools. Hence, this approach limited to gate level rather than placing transistors in the different positions in the pull-up or pull-down CMOS logic network. Thus, building a library of high-threshold and low-threshold gates and placing those gates in the performance critical and non-critical paths is more beneficial in all aspects rather than placing individual devices in the each logic gate. To achieve this, many algorithms has been proposed [Kao00] to optimally introduce low-threshold and high-threshold gates in the performance determining paths. However, the placement of low-threshold gates are limited since it contributes higher sub-threshold leakage power. Therefore, designers have to be ready to sacrifice power for high performance with low-$V_{th}$ gates and vice versa for high-$V_{th}$ devices. Depending on the power-performance demands of the application, the designers can perform optimal placement of multiple-threshold gates in the low-power and high-performance determining path with minimal design complexity. Using more than two threshold voltages shows marginal improvement in both leakage and performance advantages than the dual-threshold voltage technique. Thus, the dual-threshold voltage technique is more promising in reducing the active leakage power without much performance penalty which could be considered for energy efficient applications.

2.3.2 Sleep Transistor or Power Gating Technique

The power gating or sleep transistor is the most commonly employed technique for standby leakage power reduction in applications where standby to active time ratio is high. This can be easily achieved by introducing either nMOS or pMOS or both in series with the transistors of each logic block which creates a virtual power supply as depicted in Fig. (2.5). In practice, it is enough to have either pMOS or nMOS sleep transistor, while an nMOS sleep transistor is highly desirable because of lower ON resistance. It consists of two different operating modes (1) active mode during which the sleep transistor (ST=1) is turned on and the logic performs normal circuit operation and (2) standby mode during which the sleep transistor is completely turned off (ST=0) and the logic circuit is virtually cut-off from the power supply.
There are several factors which actually determine the optimal trade-off between the standby leakage power savings and the performance penalty due to serially stacked high-threshold power gating devices. More importantly, the selection of sleep transistor sizing and its threshold voltage affects the performance of power gating circuits. Over sizing of sleep transistor increases the potential drop in the virtual power supplies which makes adverse effect in the performance. To guarantee the proper functionality of the circuit, the sleep transistor has to be carefully sized to reduce the voltage drop across it. However, the larger sleep transistor increases both the area overhead and the switching power consumed for turning this device ON and OFF. One possible way of finding the optimal transistor sizing is based on the input vector pattern to find the worst case delay for the particular circuit. However, this may not be a suitable method for large scale integrated circuits because of the very high circuit simulation overhead. Further, an approach [Mut95] has been proposed to obtain the optimal sleep transistor width based on the mutual exclusion principle.

Alternatively, body biasing the sleep transistor (Fig. 2.5(d)) can be used to over-
come the performance degradation caused by high-threshold sleep device during active mode. Applying forward body bias lowers the threshold voltage of the sleep transistor and increases the overall performance of the logic circuits. Moreover, bringing the sleep transistor threshold voltage to its native state by withdrawing the forward body bias allows the circuit to operate in low standby power. Thus, this approach increases performance with extra area overhead, which is similar to the requirement of triple well process technology for nMOS body bias approach.

Since the suppression of standby leakage power becomes attractive and necessary for low power design, the design of leakage aware dedicated blocks becomes a challenging task to achieve low power consumption for applications with a high ratio of standby to active time. Hence, combining the effective leakage mitigation techniques with the available static logic circuits is crucial for designing energy aware dedicated blocks for power constrained wireless sensor systems. In the following section, we present the feasibility of dual-threshold voltage and power gating techniques for customized dedicated blocks using different static logic circuit families. We consider both the static and pass transistor logic circuits which are compared in terms of leakage power and performance for the above-mentioned techniques. In order to achieve maximum standby leakage power reduction in pass transistor logic families, we analyze the effect of pMOS power gating with the source signal gating. Throughout our analysis, we considered pMOS sleep transistor, because this does not require triple well process technology as needed in case of body biasing to vary the threshold-voltage depending on the mode of operation.

To summarize, a scheme which combines suitable logic circuits with leakage reduction method to optimally balance the standby leakage power and performance could be the designer’s choice. The remaining section discusses the effectiveness of above-mentioned techniques using state-of-the art logic families at different levels of abstractions (at the circuit and gate-level design hierarchy). In the following sections, we will assume the terminology like Low Leakage (LL), High Leakage (HL) and High Leakage with Header (HLH) which are used to define high-threshold, low-threshold and low-threshold with pMOS header gates respectively.
2.4 Standby Leakage Power Performance Optimization - Circuit Level

As discussed earlier, the application specific dedicated blocks with power gating would be highly beneficial for portable application, where each sub-block is controlled separately during active and standby mode of operation. Such system leverages dramatic energy savings to extend the life-time of the power source. To be more specific, we consider two different dedicated functional blocks (A and B) which consist of conventional static CMOS and pass transistor logic circuits with pMOS power gating as illustrated in Fig. (2.6). When block 'A' is active and block 'B' is in standby mode, then non-gated signals from block 'A' contribute more standby leakage power in block 'B'. However, when both of the blocks 'A' and 'B' are in standby mode, there is no need for source signal gating at block 'B' because it has intrinsic signal gating from the previous block. This problem requires extensive study of pass transistor logic circuits with power gating and source signal gating, which will be discussed in the following sections. Numerous research shows that the static CMOS has best power-delay-product compared to other logic styles [Zim97, Kos03]. Indeed other logic families such as single rail pass transistor logic (SRPL) and differential cascode voltage with pass-gate logic (DCVSPG) [Hwa97] are also a promising alternative compared to SCMOS. Hence, for our purpose we consider SCMOS, SRPL and DCVSPG circuit families with basic NAND, NOR and XOR logic.

![Figure 2.6: Dedicated LSI Blocks with pMOS Header](image-url)
functions, to study the standby leakage power and performance of dual-threshold voltage and pMOS power gating technique. Throughout our study, we considered pMOS sleep header because it require the conventional n-well process technology to perform body biasing (Fig. 2.5(4)). On the other hand, the nMOS sleep footer needs expensive triple-well process technology.

2.4.1 Static CMOS

In conventional static logic circuits (SCMOS)[Kos03], the worst case leakage current occurs when the entire pull-up devices are turned "off" (NAND gate), this significant amount of leakage current is drastically reduced by inserting high-threshold pMOS header in series with logic gate as shown in Fig. (2.7(a)). During standby mode, the switching off the logic from the power supply with header reduces $V_{ds}$ of pull up networks, which significantly reduces the leakage current at the expense of minimal performance loss. With generic power gating approach, the voltage at the virtual supply node influences the performance because of reduced $V_{gs}$ across the pull-up or pull-down transistors and $V_{th}$ increase due to body effect. By using either pMOS or nMOS sleep transistor, the designers could minimize the adverse effect introduced by virtual node voltages in the power gating technique. Therefore, its more desirable to have one sleep transistor rather than using conventional power gating technique.

2.4.2 Single Rail Pass Transistor Logic

In Single Rail Pass Transistor Logic (SRPL) [Zim97], an nMOS (pMOS) pass transistor with the signal input fed to the drain(source) and the signal output taken from source(drain). The propagation of signals in the pass network in controlled by the gate inputs. Here, the logic functionality is defined by the HL transistors with LL pMOS header as illustrated in Fig. 2.7((b)&(c)). Unlike SCMOS, the SRPL logic gate creates a path from the source terminal to the output node based on the control gate input. Hence, the leakage current in nMOS pass transistor logic gates mainly depends on pass signal inputs and the logic-defined. Thus, these leakage currents flow either through the input or output static buffers, which establish the path from supply to ground. The leakage current flowing either from the inverting or non-inverting logic, input and output static buffers are shown by soft dotted arrow
Figure 2.7: (a) Conventional static CMOS NAND with pMOS Header (b) & (c) HL-SRPL NAND & AND gates with output level restorer (d) HL-DCVSPG NAND/AND pMOS Header

lines as shown in Fig. (2.7)((b)&(c)).

In the non-inverting pass logic (Fig. 2.7(c)), the chosen input pattern $A=0, B=1$ causes the leakage current to flow and these current charges the intermediate node 'X', which increases the leakage current nMOS "off" transistor. It shows that the input vector "01" or "10" contributes more leakage, when off current (sub-threshold leakage current) from the logic charges the node 'X'. On the other hand, the input
"00" or "11" contributes to lower leakage current compared to other input vectors. Therefore, the leakage current flows mainly from the input and output buffers. In case of inverting pass logic (Fig. 2.7(b)), the static buffer at each input source signal is gated by the header, which reduces leakage current flowing from the OFF transistors. Thus, the input pattern "00" or "11" has no contribution to total leakage current and the patterns "01" or "10" reduce leakage current due to implicit source signal gating.

Therefore, the non-inverting pass logic with header has shown less effect on leakage reduction of two reasons. At first, there is no input signal gating at the source input to reduce the drain source potential during standby mode. Secondly, the pMOS header at the output level restoration has not taken part in reducing off currents from the nMOS pass logic. This implies that the non-inverting pass logic requires extra buffer gating at its source end to reduce the sub-threshold leakage current, whereas inverting pass logic function reduces leakage power with an implicit embedded static inverter.

### 2.4.3 Differential-Cascode Voltage Switch with Pass-Gate Logic

The ratioless circuit technique is designed using a pass-gate logic tree instead of nMOS logic tree in the conventional DCVS circuit, which eliminates the floating node problem [Hwa97]. In Differential-Cascode Voltage Switch with Pass-Gate Logic (DCVSPG), the inverting and non-inverting logic block consists of HL transistors and power gated pMOS header for both logic blocks and input buffers as shown in the Fig. 2.7(d). Similar to SRPL, DCVSPG gates requires additional signal gating at all its non-inverting source inputs to attain maximum leakage current reduction. From the chosen input pattern A=0 and B=1, the leakage current flows from higher to the lower potential node (\texttt{gnd}) either through the logic or input buffers as shown in the soft dotted arrow lines in Fig. 2.5((e)&(f)). It shows that the DCVSPG HLH gates has embedded buffer gating for inverting source signals, however an extra gated buffer should be introduced at the non-inverting source signal to achieve maximum leakage current reduction.
2.4.4 Simulation Results and Discussion

The simulations are carried out in cadence spectre simulator using 180nm mixed mode process technology. To compare the different logic styles using various leakage reduction techniques under identical conditions, a load capacitance of 1.5fF is taken throughout our simulation. The power-delay-product is computed from the product of worst case delay and switching power of the logic gates.

Figure (2.8) illustrates the simulated values of PDP for HL, HLH and LL NAND, NOR and XOR gates using SCMOS, SRPL and DCVSPG logic styles respectively. The power-delay-product of NAND and XOR High Leakage Header shows 24% and 27% increase than HL gates with 99.9% reduction of standby leakage power. This is due to two main reasons (1) the serially connected sleep transistor increases the circuit delay minimally which in turn increases the PDP and (2) the active sub-threshold leakage current from the low-threshold logic transistors also increases its PDP metric. In the case of SRPL gates, we observed that the PDP of HLH NAND and XOR with buffer have increased by 2.3 and 3.4 times than without

Figure 2.8: power-delay-product of HL, HLH and LL gates using SCMOS, SRPL and DCVSPG logic circuits
buffer. However, it achieves higher standby leakage power reduction due to the buffer gating.

The power-delay-product of DCVSPG NAND and XOR gate with buffer shows 0.6 and 0.7 times increase compared to HLH gates. Thus, the DCVSPG with buffer shows greater advantage in terms of PDP than SRPL gates. In standby mode of operation, the worst case leakage power has computed for HL, HLH, LL and HLH with buffer gates are shown in the Table (2.1). The SRPL NAND and NOR gates has not shown greater impact on leakage current reduction due to the lack of pass signal gating. The pass transistor XOR HLH gate shows similar characteristics of SCMOS due to the embedded buffer gating and it attains the maximum leakage power reduction. The SRPL and DCVSPG gate reduces standby leakage power like SCMOS at the expense of extra transistor count for source signal gating. This clearly suggests that, the signal coming from the previous stages to the source input of pass logic gates should be gated carefully when the dedicated blocks (pass-logic) switch between the active and standby modes.

Table 2.1: **Standby-leakage power of SCMOS, SRPL & DCVSPG logic circuits**

<table>
<thead>
<tr>
<th>Logic styles</th>
<th>Leakage techniques</th>
<th>NAND(nW)</th>
<th>NOR(nW)</th>
<th>XOR(nW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCMOS</td>
<td>HL</td>
<td>49.35</td>
<td>97.73</td>
<td>132.7</td>
</tr>
<tr>
<td></td>
<td>HLH</td>
<td>0.002</td>
<td>0.010</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>LL</td>
<td>0.023</td>
<td>0.255</td>
<td>0.292</td>
</tr>
<tr>
<td>SRPL</td>
<td>HL</td>
<td>180.6</td>
<td>155.9</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td>HLH</td>
<td>48.81</td>
<td>110.3</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>HLH+Gated input</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>LL</td>
<td>0.408</td>
<td>0.396</td>
<td>0.408</td>
</tr>
<tr>
<td>DCVSPG</td>
<td>HL</td>
<td>180.5</td>
<td>204.8</td>
<td>204.8</td>
</tr>
<tr>
<td></td>
<td>HLH</td>
<td>48.71</td>
<td>20.29</td>
<td>102.7</td>
</tr>
<tr>
<td></td>
<td>HLH+Gated input</td>
<td>0.009</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>LL</td>
<td>0.408</td>
<td>0.524</td>
<td>0.524</td>
</tr>
</tbody>
</table>
2.5 Standby Leakage Power Performance Optimization - Gate Level

In this section, we made an attempt to combine the dual-threshold voltage and power-gating technique together to effectively reduce both the active and standby power for ultra-low power applications. Therefore, we introduce the High-Leakage Header (HLH) gates with in-built pMOS power gating for low-threshold logic cells which are placed in the performance determining critical paths. The basic logic gates are designed with HLH for all the above-mentioned static logic circuit families.

First, we analyzed the efficiency by placing HL (or) HLH and LL gates in the critical and non-critical paths respectively, to minimize the leakage power consumption. In addition, the HL Header with input buffer gating has been evaluated for pass logic adders.

To evaluate the dual-threshold header scheme, an adder cell in the hierarchical fashion, as shown in Fig. (2.9) is designed using Low-Leakage, High-Leakage, dual-threshold voltage with and without high-leakage Header gates respectively. In dual-threshold header, the critical path HLH gates and the non-critical path LL gate shows minimal performance loss in the active mode with higher cutback of leakage power in the standby mode compared to traditional dual-threshold techniques. As discussed in the previous sections, the significant leakage current flows when the
2.5 Standby Leakage Power Performance Optimization - Gate Level

Figure 2.10: Power-Delay-Product and standby-leakage power for SRPL and DCVSPG adder circuit

HLH gates draws non-gated signal from the previous block in the standby mode. In case of pass logic adders, the "XOR" and "AND" gates receive signal 'A', 'B' and "Ci" from the previous stages and it has not been gated by the input buffer and therefore it allows the leakage current to flow even when the header is turned off. These gates have intrinsic embedded buffer gating for inverting signals and there is no buffer gating for non-inverting signals. Like inverting signals, the non-inverting signals should be gated with additional input buffer to achieve maximum leakage power reduction. Hence, the critical path HLH gates which receive input signals from the previous stages have to pass through the gated input buffer to reduce the leakage currents in the pass transistor gates.

2.5.1 Simulations Results and Discussion

Extensive simulations have been carried out for all possible input combinations to get the worst case leakage power. In dual-threshold header, the SCMOS adder degrades the performance by a factor of 4.7% when compared to dual-threshold
without header as shown in Table (2.2). For dual-threshold header with input buffer, the power-delay-product of SRPL and DCVPSG adders are increased by a factor of 2.2% and 11.7% when compared to dual-threshold header without input buffer. In case of standby leakage power, the dual-threshold with header SRPL and DCVPSG adders without source signal gating showed less percent of leakage reduction than that of SCMOS adders. The dual-threshold adder with and without buffer gating reduces the standby leakage power by 96% and 77.7% for SRPL and DCVPSG respectively as shown in Fig. (2.10). Thus, the dual-threshold adder with buffer gating reduces the standby leakage power on an average of (95)-(99)% than the dual-threshold without header. The standby leakage power of dual-threshold header with and without buffer gated SRPL and DCVPSG shows similar property of SCMOS power gating, provided the signals from the previous blocks are perfectly gated. These additional buffers at the non-inverting inputs increase the PDP with greater reduction of standby leakage power during standby mode. It shows that the input buffer gating should be introduced at both ends of the pass transistor logic gates to effectively reduce the standby leakage power.

Table 2.2: Power-Delay-Product and standby-leakage power for different schemes using SCMOS full-adder circuit

<table>
<thead>
<tr>
<th>Different Schemes</th>
<th>$PDP(fJ)$</th>
<th>Leakage Power (nW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Leakage Adder</td>
<td>20.28</td>
<td>1.59</td>
</tr>
<tr>
<td><strong>dual-threshold voltage with Header</strong></td>
<td><strong>22.99</strong></td>
<td><strong>0.44</strong></td>
</tr>
<tr>
<td>dual-threshold voltage without Header</td>
<td>23.58</td>
<td>378</td>
</tr>
<tr>
<td>High Leakage Adder</td>
<td>24.01</td>
<td>1136</td>
</tr>
</tbody>
</table>

In this section, we addressed different sources of leakage current and their effective mitigation techniques for low power applications. We showed that, the dual-threshold voltage and power-gating techniques are highly efficient for active and standby leakage power reduction. Further, we have analyzed a co-design approach in which the pMOS gating is introduced in the dual-threshold technique to determine the optimality between the power and performance for different static logic circuits. In addition, we examined the method of introducing source signal gating at each inputs in the pass logic circuits to design power-constrained sensor node.
3 Energy Efficient Circuit Design

Energy efficient circuit design becomes increasingly crucial to meet the power and performance demands for modern day portable electronics. To reduce energy-per-cycle as well as to improve performance in the ultra-low voltage region, we discuss different forward body bias configurations in the sub-threshold, near-threshold and above-threshold regime. We begin our study towards minimum energy-per-cycle using symmetric and asymmetric forward body bias schemes by proposing several approaches to energy efficient design with performance advantages. To extend our study further, we analyze the characteristics of forward body bias in the presence of transistor width, noise margin and process variation in the ultra-low voltage regime.

3.1 Introduction

Low energy LSIs are primary concern in today’s battery enabled ubiquitous computing applications. Conventionally, the energy efficiency is improved quadratically by lowering the supply voltage at the expense of severe delay penalty. However, the supply voltage reduction by the conventional approach is limited by three main factors (1) Leakage power becomes major portion of total power consumption at lower supply voltages (2) severe performance degradation caused by fabrication process deviation, supply voltage and temperature variation becomes larger below certain supply voltages and (3) severe static noise margin at very low supply voltages [Nar04].

As process technology scales down, the ITRS roadmap predicts that supply voltage is approaching towards sub-1V operation for higher energy efficiency by lowering the threshold voltage as shown in Fig. (3.1). It shows that operating supply voltage for Low Operating Power (LOP) devices is likely to scale down from 0.8V (65nm) to
0.5V at 22nm node. In general, there are two major approaches which targets either low energy (power) or higher energy efficiency by meeting performance requirements in the low-$V_{dd}$ region, (1) operating at sub-threshold region ($V_{dd} < V_{th}$) when performance is not a major concern targets only ultra-low power low performance solution [Cal04] and (2) operating at ultra-low $V_{dd}$ region achieving moderate performance improvement using forward body bias (FBB) [Arn04]. For instance, approach (1) is most likely suitable for applications like wireless sensor nodes [Jay05a] in which the minimum power consumption is the primary concern over performance design metric to extend the battery lifetime. On the other hand, approach (2) which provides flexible performance improvement at the expense of minimal power consumption which is best suitable for wide-range of applications like bio-medical, wireless surveillance system and etc.

Therefore, it is clear that approach (1) mainly targets the minimum energy-per-cycle with significant circuit delay penalty. However, the higher impact on process variation because of exponential dependency on the threshold voltage with the corresponding ultra-low performance solution circumscribed the effectiveness of sub-threshold operation in the nanometer regime. It imposes several challenges for circuit designers to design successful sub-threshold circuits. In order to study these issues, the effective threshold voltage tuning method (approach (2)) in the ultra-low voltage region using triple-well process technology is considered in this chapter.

Figure 3.1: Technology scaling according to ITRS roadmap predication
Conventionally, the method of applying forward body bias symmetrically to the entire CMOS circuit enhances performance by lowering the threshold voltage, which is widely used for high-performance applications [Nar03]. In the active-mode, the forward bias voltage is applied to both pull-up and pull-down devices to improve the driving current capability. In case of standby-mode, the reverse body bias raises the threshold voltage which lowers the standby leakage power. In the following section, we analyze the characteristics of energy efficient LSIs with forward body bias in the sub-threshold, near-threshold and above-threshold region to determine the optimum between energy and performance design metric. Furthermore, we investigate the characteristics of symmetric and asymmetric forward body bias configurations in all the three regions. Since our focus is on ultra-low voltage applications, we consider the supply voltage of preferably near 0.5V $V_{dd}$ with 13-stage adder based delay chain throughout our study.

The rest of the chapter includes the theoretical minimum supply voltage ($V_{dd_{min}}$) and the corresponding energy-per-cycle ($E_{min}$), the figure of merit for highly energy efficient design, the low-leakage and high-speed device characteristics with FBB, the symmetric and asymmetric FBB configuration and further, it discusses the critical circuit design parameters like noise margin and transistor-width. Finally, the mitigation of process parameter variations with respect to forward body bias in the low-$V_{dd}$ regime is discussed.

### 3.2 Theoretical $V_{dd_{min}}$ and $E_{min}$

This section discusses the theoretical lower limit of operating supply voltage and the corresponding energy-per-cycle that is required to perform fully functional CMOS logic in the ultra-low $V_{dd}$ space. The minimum allowable operating supply voltage was derived by [Mei00] as

$$V_{dd_{min}} = 2V_T \left[ 1 + \frac{C_{fs}}{C_{ox} + C_d} \right] \ln \left( 2 + \frac{C_d}{C_{ox}} \right)$$  \hspace{1cm} (3.1)

where $C_{fs}$ is the fast surface state capacitance per unit area, $C_{ox}$ is the gate-oxide capacitance per unit area, $C_d$ is the channel depletion region capacitance per unit area and $V_T = (kT/q) = 25.875$ mV for 300°K.
The above-mentioned $V_{dd_{min}}$ depends on the subthreshold slope factor, which can be expressed as follows

$$S = \left( \frac{d(\log_{10} I_{ds})}{dV_g} \right)^{-1} = 2.3V_T \left( 1 + \frac{C_d}{C_{ox}} \right)$$  \hspace{1cm} (3.2)$$

where $(1 + \frac{C_d}{C_{ox}})$ is called the body-effect coefficient and typically lies between 1.1 and 1.4 [Tau98]. To rewrite the equation (3.1) in terms of sub-threshold slope, the equation (3.2) in (3.1) can be expressed as

$$V_{dd_{min}} = 2V_T \left[ 1 + \frac{C_{fs}}{C_{ox} + C_d} \right] \ln \left( 1 + \frac{S}{2.3V_T} \right)$$  \hspace{1cm} (3.3)$$

For an ideal MOSFET (i.e., a device with a sub-threshold slope of 60mv/decade at 300°K) $C_{fs} << C_{ox}$ and $C_d << C_{ox}$, thus the minimum allowable supply voltage can be written as

$$V_{dd_{min}} = 2V_T \ln \left( 1 + \frac{S}{2.3V_T} \right)$$  \hspace{1cm} (3.4)$$

The above equation represents the theoretical minimum allowable supply voltage with respect to temperature and sub-threshold slope factor, which is shown in Fig. 3.2(a). It shows that, the $V_{dd_{min}}$ increases with the ambient temperature and the sub-threshold slope factor. In case of ideal MOSFET, the lowest allowable supply voltage can be written as

$$V_{dd_{min}} = 2V_T \ln(2) = 35.87mV$$  \hspace{1cm} (3.5)$$

The industrial 130nm technology is used in our simulation, a $V_{dd_{min}} = 47.67mV$ for the minimum size device with sub-threshold slope of nearly 90mV/decade at 300°K is computed using equation(3.4). It shows that, the theoretical minimum supply voltage allows circuits to operate into the deep sub-threshold regime provided the threshold voltage are much larger than $V_{dd_{min}}$. In order to analyze this limit, the spice simulation is carried out for minimum size 31-stage inverter delay chain which oscillates at $V_{dd_{min}}$ with the reduced internal voltage swing (less than 30mV). This is due to inadequate current drive at 47.67mV is not sufficient to switch the intrinsic capacitance to make full voltage swing. It clearly illustrates that, it is highly impossible and not desirable to operate the practical circuits at theoretical $V_{dd_{min}}$ due to many practical reasons such as higher susceptibility to noise and process
3.2 Theoretical $V_{dd\min}$ and $E_{min}$

![Graph showing theoretical minimum allowable supply voltage and minimum switching energy per cycle with respect to temperature and sub-threshold slope factor.](image)

Figure 3.2: (a) Theoretical minimum allowable supply voltage with respect to temperature and sub-threshold slope factor (b) Minimum switching energy-per-cycle exacerbates with temperature fluctuations.

In addition to theoretical $V_{dd\min}$, Swanson and Meindl illustrated the fundamental limit of minimum switching energy of MOSFET at $V_{dd\min}$ by assuming the minimum possible electronic charge (single electron) stored on the gate capacitance. In practice, the switching energy of MOSFET can be written as

$$E_s = \frac{1}{2} Q_c V_{dd}$$  \hspace{1cm} (3.6)

Where $Q_c$ is the amount of charge stored on the capacitance and $V_{dd}$ is the operating supply voltage. Equation (3.4) and (3.6) shows the fundamental limit of minimum possible operating voltage and the corresponding switching energy of single MOSFET. To derive the lowest possible switching energy with the theoretical $V_{dd\min}$, the equation (3.4) is rewritten with equation(3.6) as

$$E_{s\min} = \frac{1}{2} Q_c V_{dd\min} = Q_c V_T \ln \left( 1 + \frac{S}{2.3 V_T} \right)$$  \hspace{1cm} (3.7)

The above equation suggests that the switching energy entirely depends on the capacitance and the operating voltage. In case of ideal MOSFET with $Q_c = q$, 39
then the lowest possible single MOSFET switching energy \( E_{s\text{min}} = KT\ln(2) \) which is equivalent to \( 2.869 \times 10^{-21} \) Joules at 300 \( ^\circ \)K. Further, the ambient temperature dependency on the switching energy-per-cycle is shown in Fig. 3.2(b), which clearly indicates that the switching energy exacerbates with temperature. In case of static CMOS inverter with pull-up width equal to three times of pull-down width, the \( E_{s\text{min}} \) is calculated to be four-fold that of single MOSFET. In other words, \( Q_c \) is actually \( C_{\text{ox}}V_{dd} \) in a static CMOS inverter, where \( C_{\text{ox}} \) is the sum of n and p-MOSFETs.

### 3.3 Energy and Performance - Figure of Merit

Like supply voltage, the threshold voltage is one of the key tuning parameter in the CMOS VLSI design. The total power consumption and performance can be estimated based on three primary factors namely threshold voltage (\( V_{th} \)), supply voltage (\( V_{dd} \)) and the operating clock frequency (\( f_{clk} \)). Thus, the total power consumption can be written as [Cha92],

\[
P_T = P_{AC} + P_{ST}
\]

Where \( P_{AC} \) represents active switching power and \( P_{ST} \) is the static leakage power. Since we considered the delay chain throughout our analysis, the average total power consumption can be written as,

\[
P_T = \frac{1}{2} n. (\alpha C_s) V_{dd}^2 f + (n.I_{\text{Static}}) V_{dd}
\]

(3.8)

Where \( \alpha \) represents the activity factor (the average number of transitions per clock cycle), 'n' is the number of logic stages or logic depth, \( C_s \) represents the switching capacitance, 'f' represents delay chain frequency and \( I_{\text{Static}} \) represents static leakage current. As one can see in the above equation, the short-circuit power can be ignored safely because in the typical assumption that the direct-path current approaches zero when the supply voltage reduces below \( V_{th,n-FET} + |V_{th,p-FET}| \) and it becomes the entirely sub-threshold leakage current below this voltage [Han06]. Figure (3.3)(a) shows that, the active power is dominant over the static power till certain supply voltage (0.2V), below which the static power is increased dramatically due to larger logic depth. The 13-stage adder oscillates with full output voltage swing until 0.15V and below which the functionality collapses.
### 3.3 Analytical Study of Minimum Energy-Per-Cycle

Since our focus is on energy efficient design, we consider the energy-per-cycle or power-delay-product and energy-delay-product as the primary figure of merit to analyze energy efficiency and the corresponding performance in the low-$V_{dd}$ regime. Like equation (3.8), the energy/cycle for an n-stage delay chain can be written as,

$$E_T = E_{AC} + E_{ST}$$

$$E_T = \frac{1}{2} n. (\alpha C_s) V_{dd}^2 + (n. I_{Static}) V_{dd} n. t_d$$

(3.9)

The above equation represents the total average energy-per-cycle which is calculated by summing up both the active and static energy of n-stage delay chain. It can be seen that the active energy has quadratic dependence, while the static leakage energy has linear dependence on the operating supply voltage. It further shows that lowering the supply voltage reduces the active energy, whereas the static energy exacerbates due to larger circuit delay in the low-$V_{dd}$ regime.

To extend our study further, we focused on determining the circuit switching delay for short-channel devices. However, the accurate CMOS circuit delay '$t_d$' is quite complex when the non-linear CMOS gate is taken into account in the nanometer regime, therefore we have taken the simple first order derivation which is shown below,

$$t_d = \frac{(\alpha C_s) V_{dd}}{I_{on}}$$

(3.10)

Here, $I_{on}$ represents the sub-threshold current when the supply voltage is lower than the threshold voltage, and it can be expressed as follows,

$$I_{on} = \frac{W}{L} \mu_{eff} C_{ox} (m - 1) (V_T)^2 e^{(V_{gs} - V_{th})/mV_T} \left(1 - e^{-V_{ds}/V_T}\right)$$

(3.11)

where $V_T = (kT/q)$ called as thermal voltage, 'T' represents temperature, 'k' is Boltzmann’s constant, 'q' is the electronic charge, $\mu_{eff}$ is the effective mobility, $C_{ox}$ is the oxide capacitance per unit area, 'W' is the gate width and 'm' is the sub-threshold slope factor. Equation (3.11) shows the sub-threshold current dependency on the threshold voltage and the effective channel length. The threshold voltage variation is highly sensitive to $L_{eff}$ and drain induced barrier lowering (DIBL) which actually fluctuates the sub-threshold current significantly. By replacing $\alpha C_s = C_{eff}$ and
Figure 3.3: (a) Power consumption for a 13-stage adder decreases with supply voltage (b) The optimal minimum energy-per-cycle and the corresponding supply voltage (130nm technology)

substituting equation (3.10) in (3.9), the final expression for the overall energy-per-cycle can be rewritten as,

$$E_T = \frac{1}{2} n C_{eff} V_{dd}^2 + (n I_{stat}) V_{dd} \left( n \frac{C_{eff} V_{dd}}{I_{on}} \right)$$ (3.12)

$$E_T = \frac{1}{2} n^2 C_{eff} V_{dd}^2 \left[ \frac{1}{n} + \frac{I_{stat}}{I_{on}} \right]$$ (3.13)

It can be seen that the total energy/cycle varies with the ratio of static leakage current to the active switching current. This ratio will change depending on the region of operation, for example in the above-threshold region, the on-current is significant over the static current and vice versa in the sub-threshold region. In order to study the overall energy/cycle in different region of operation, we consider three major cases like case(1) when $I_{on} >> I_{stat}$ i.e., in the above-threshold region ($V_{dd} > V_{th}$), case(2) the below threshold-region $I_{on} = I_{stat}$ = minimum operating current i.e., in the operating region where the on-current becomes equivalent to the sub-threshold current (both are same in this case) and finally case(3) the below-threshold ($V_{dd} < V_{th}$) (or) below minimum operating region. Lets first take case(1),
when $V_{dd} > V_{th}$ and correspondingly $I_{on} >> I_{static}$ has dominant active energy over the static energy in the overall energy/cycle. Case(2) shows that the lowest possible driving current in the deep-sub-threshold region when $V_{gs} < V_{th}$, i.e. $I_{Static} = I_{on}$, $rac{I_{Static}}{I_{on}} = 1$ and therefore the corresponding energy-per-cycle can be written as,

$$E_{min} = \frac{1}{2} n C_{eff} V_{dd}^2 (1 + n)$$  \hspace{1cm} (3.14)$$

The above equation represents the minimum energy/cycle and it occurs when both $I_{on}$ and $I_{Static}$ are same at certain supply voltage. This corresponding voltage called minimum operating supply voltage ($V_{dd_{min}}$). In case (3), the static energy is dominant because of significant leakage over active energy due to larger ($\frac{I_{Static}}{I_{on}}$) ratio in the sub-threshold region. Moreover, higher the logic depth and larger will the static energy, which increases the overall energy/cycle since it has quadratic dependency on the number of logic stages.

The insertion of circuit delay in Equation (3.8) gives the average energy-per-cycle which is given in Equation (3.9). In this case, the energy-per-cycle means the amount of work done in a single clock period rather than single clock cycle/transition. As we know, the circuit delay increases significantly especially when it enters into the sub-threshold region, because of exponential dependency on sub-threshold current. Though, the sub-threshold current reduces with lower supply voltage, the increase in delay is so rapid that the static leakage energy quickly overtakes the active energy in the deep sub-threshold region. Figure (3.3) illustrates the average power consumption and the energy-per-cycle of 13-stage adder delay chain in an industrial 130nm CMOS technology. It shows that average energy-per-cycle decreases with supply voltage until it reaches the minimum operating voltage ($I_{on} = I_{Static} =$ minimum allowable operating current). In this case, the supply voltage is scaled down and the threshold voltage is fixed at approximately 0.38V. For the circuit under consideration, the optimal minimum energy/cycle with the corresponding supply voltage occurs in the sub-threshold region and it yields significant energy reduction. This results is confirmed with several publications [Cal04] and [Han06], which found sub-threshold voltages to be optimal for an inverter chain and an FIR filter with fixed threshold voltage respectively. On the other hand, [Wan02] reports the energy saving by simultaneously scaling both $V_{dd}$ and $V_{th}$ in the sub-threshold region for
higher energy efficiency when performance is not of major concern. However, the low-performance solution and large variation due to $L_{eff}$ and DIBL reduces the effectiveness of optimal sub-threshold operation for ultra-low power applications. For example, reasonable $6\sigma$ 100-mv $V_{th}$ mismatch disturbs MOSFET current ratio by 1.17x in the above-threshold region, whereas it upsets the current matching by greater than 10x in sub-threshold operation. Thus, the following section analyzes the threshold voltage with forward body bias to improve the energy efficiency by compensating the fluctuation introduced by manufacturing device parameters.

### 3.3.2 Threshold Voltage Analysis with Body Bias and Temperature

As we discussed in the previous equations (3.11) and (3.14), the threshold voltage and the temperature plays a vital role in determining the minimum energy/cycle and the optimal minimum operating voltage in the sub-threshold region. The threshold voltage $V_{th}$ of a MOSFET can be written as,

$$V_{th} = V_{th0} + \gamma(\sqrt{|2\phi_F| - V_{bs}} - \sqrt{|2\phi_F|})$$  \hspace{1cm} (3.15)

where

$$\Delta V_t = \gamma(\sqrt{|2\phi_F| - V_{bs}} - \sqrt{|2\phi_F|})$$  \hspace{1cm} (3.16)

and

$$\gamma = \frac{\sqrt{2\varepsilon_{si}qN_a}}{C_{ox}} = \sqrt{2\varepsilon_{si}qN_a}\frac{t_{ox}}{\epsilon_{ox}}$$  \hspace{1cm} (3.17)

Where $V_{bs}$ is the body to source voltage, $\phi_F$ is the fermi potential, $t_{ox}$ is the gate-oxide thickness, $\epsilon_{ox}$ is the dielectric constant of $SiO_2$, $\epsilon_{si}$ is the permittivity of silicon and $N_A$ is the carrier concentration in intrinsic silicon. The constant $\gamma$, characterizes the body effect, and is called the body-bias coefficient. The above equation shows that, as the process technology scales into sub-100nm regime, the body-effect decreases because of smaller gate oxide thickness ($t_{ox}$) and lower channel implants ($N_a$). In recent years, the purpose of body biasing has been extending from the standby sub-threshold leakage reduction to the active power reduction by lowering the threshold voltage. Further, it helps in compensating variations in $V_{th}$ in high-end microprocessors, and the body bias range has been extended from reverse body.
bias (RBB) to forward body bias (FBB) in order to maintain the coverage of $V_{th}$ change in scaled CMOS devices. In the next section, we discuss the low-leakage (LL) and high-speed (HS) device characteristics with respect to forward body bias in the ultra-low $V_{dd}$ regime.

(I). Characteristics of Forward Body Bias Device

As we discussed in the previous section, ITRS roadmap predicts that the nominal supply voltage is relatively high and this will slowly scale down to low standby power (LSTP) requirements, because the saturation threshold voltage is kept high to maintain the sub-threshold leakage current very low. On the other hand, for low operating power (LOP) requirements $V_{dd}$ scales down rapidly in order to keep the dynamic power consumption to a lowest value. It predicts that LOP supply voltage will be scaled below 0.6V by the year 2015, which is quite interesting for us to analyze the forward body bias in the ultra-low $V_{dd}$ region as well.

![Figure 3.4](image.png)

Figure 3.4: High-speed MOSFET threshold voltage reduction and the corresponding bulk-current increase for different body bias voltage at elevated temperature (pMOS)

From Equation (3.15), $\gamma$ is lowered in a short channel transistor, because the channel potential is influenced more by the drain voltage than the body bias because...
of the DIBL effect. Moreover, the widening of drain-substrate depletion layer with reverse body bias (RBB) worsens the short-channel effects (SCE), which varies $V_{th}$ across a die. Thus, the SCE and DIBL are particularly severe in low-$V_{th}$ devices because reducing the channel doping to lower $V_{th}$ causes the channel depletion depth to become larger. One can see that, the range of body bias is extended from the reverse body bias to forward body bias and thus it can be applied to high-$V_{th}$ to bring $V_{th}$ down to the target value.

To insure that the parasitic junctions are not heavily turned on, the maximum body bias is selected to be 0.5V and above which the body-to-source diode current increases rapidly as shown in Fig. (3.4). A 0.5V FBB is practical since no latch up effect was observed by the forward bias up to a 4-V supply voltage [Oow98]. Conventionally, $V_{bsn}$ and $V_{bsp}$ are connected to $V_{ss}$ and $V_{dd}$ respectively, without any body bias control. When FBB is applied, the static-leakage current due to $V_{th}$ reduction, p-n junction leakage currents and the parasitic bipolar current starts rising, but this could be limited by the amount of body bias applied and the operating supply voltage. Figure (3.5) shows the cross section of triple-well CMOS structure in which pMOS and nMOS devices are placed in n-well and p-well regions respectively. To avoid the parasitic bipolar current, a thick n-isolation layer (deep n-well) is placed underneath the p-well and n-well regions which reduces the bipolar current.
amplification factor considerably [Miy00]. On the other hand, when the supply voltage scales below 0.5V ($V_{bs} = 0.5V$), the voltage across the p-well and n-well diode becomes positive which forward biases the well-diode.

Table (3.1) represents the different modes of parasitic diode operation when the supply voltage is larger than the body bias, equal to body bias and smaller than the body bias. For higher supply voltages, most of the parasitics are reverse biased whereas it becomes forward biased in the lower supply voltage. However, when both $V_{dd}$ and $V_{bs}$ are same, then the well diode and body-to-drain diode (low-to-high transition) operates in zero bias mode. Hence, it is quite safe to operate when the supply voltage is either higher or equal to the applied FBB. On the other hand, operating at lower supply voltage has superior performance over the higher supply voltage. Though the well diode forward biased in the lower supply voltage, the performance improvement and the minimal parasitic currents is still negligible compared to higher supply voltage.

To study the device characteristics, the minimum size low-leakage ($LL$) and high-speed ($HS$) nMOS devices from industrial 130nm triple well process ($W_n=0.16\mu m$, $L_{min}=0.12\mu m$) are considered. Figure (3.6) shows that the forward body bias is highly efficient for $LL$ compared to $HS$ devices due to larger body-effect of high-threshold device. Further, it shows that the drive current of $LL$ device always

Figure 3.6: MOSFET threshold voltage and the corresponding drive-current for low leakage and high-speed device with wide-range of body bias voltage

3.3 Energy and Performance - Figure of Merit
improving with FBB even beyond diode cut-in voltage (0.6V) but it’s not the case for HS devices. In both HS and LL devices, the maximum allowable 0.5V FBB reduces the threshold voltage by nearly 11% and 12% in the sub-threshold and above-threshold region which is shown in Table(3.2). On the other hand, the drain current improvement by 0.5V FBB in sub-threshold region ($V_{dd} = 0.2V$) is larger than that of above-threshold region (0.6V) in both HS and LL devices. Though it has considerable $V_{th}$ reduction irrespective of operating region, the improvement in drive current diminishes at higher supply voltage because of dominant DIBL effect. Thus, the reduction in threshold voltage due to DIBL in the above-threshold region makes it difficult to find good compromise between the applied FBB and the drain-current improvement.

Moreover, the $I_{off}$ increases dramatically with FBB compared to the $I_{on}$ improvement, which is independent of operating region and mode of $V_{th}$ device chosen. This is due to the extra diode junction leakage current introduced by FBB, which
3.3 Energy and Performance - Figure of Merit

Figure 3.7: Positive temperature dependence of Low-Leakage and High-speed devices in low $V_{dd}$ region at elevated temperature (between $25^\circ C$ to $125^\circ C$)

increases the $I_{off}$ considerably in the low-$V_{dd}$ regime. FBB has very strong dependency on temperature, the $V_{th}$ thus the corresponding $I_{on}$ and $I_{off}$ increase at elevated temperatures which will be discussed in detail in the following sections.

(II). Impact of Temperature Dependence

As the supply voltage scales below zero-temperature coefficient point [Sak01] in the nanometer MOSFET, the positive temperature dependence on drain current is observed in the ultra-low $V_{dd}$ region [Bry01]. The drain current dependence on temperature can be written as,

$$I_{ds} = \mu(T)(V_{dd} - V_{th}(T))^\alpha$$ (3.18)

$$V_{th}(T) = V_{th}(T_0) - \kappa(T - T_0)$$

$$\mu(T) = \mu(T_0)(\frac{T}{T_0})^{-m}$$

where $V_{th}(T)$ and $\mu(T)$ represents the threshold voltage and mobility factor dependence on the temperature, $T$ means temperature and $T_0$ represents room temperature, $\kappa$ represents the threshold voltage temperature coefficient ($2.5mV/K$) and $m$ represents the mobility temperature exponent (1.5). Equation (3.18) represents the drain current which varies depending on the operating temperature and the supply.
voltage. It is well known that both the mobility and threshold voltage reduces with the increase in temperature which in turn degrades the drain current in high $V_{dd}$ region. As the supply voltage approaches the threshold voltage in the ultra-low $V_{dd}$ region, the reverse temperature phenomena occurs (positive temperature dependency) and this is mainly due to the dominant threshold voltage variations. The threshold voltage reduction with increase in temperature improves both HS and LL device drain current as shown in Fig. (3.7).

Equation (3.11) shows the exponential dependence of sub-threshold current on $V_{th}$ and thermal voltage ($V_T$). Thus, the $V_{th}$ is lowered with both FBB and higher temperature, which increases the drain current significantly in the ultra low-voltage region. For example, at higher temperature 125°C with 0.5V FBB, the drain current of HS and LL device improved by 10X and 132X compared to 25°C without any FBB condition. It shows that sub-threshold current is more sensitive to both temperature and FBB and it increases dramatically with temperature. Further, the FBB characteristics are strongly dependent on the device characteristics and operating temperature in the ultra-low $V_{dd}$ region.

3.4 Multi-stage Adder based Delay Chain - Circuit Under Study

To evaluate the different FBB configuration in the ultra-low voltage region, we consider the 13-stage adder based delay chain as shown in Fig. (3.8). It consists of $n$ stages of conventional static CMOS adder blocks with extra circuitry at the input stage to perform oscillation. Figure (3.9) shows the schematic cross-section of minimum size adder with the corresponding input ($A, B, Cin$) and output ($SUM$ and $CARRY$) signals. In this delay chain, each adder block consists of common input which tied A and B signals together, carry input and the sum and carry outputs. Each adder block is cascaded in such a way that both sum and carry output is rippled to the next stage to perform oscillation.

There are two cases in which the adder oscillates,

Case(1) When $A = B = \text{HIGH}$ and $Ci = \text{LOW}$, then $So = \text{LOW}$ and $Co = \text{HIGH}$ and Case(2) When $A = B = \text{LOW}$ and $Ci = \text{HIGH}$, then $So = \text{HIGH}$
3.4 Multi-stage Adder based Delay Chain - Circuit Under Study

Figure 3.8: The 13-Stage adder based delay chain with an enabling input block to trigger oscillation

and Co = LOW. Therefore, the sum and carry function of a conventional adder can be written as,

\[
So = A \oplus B \oplus Ci = Ci.(A.B + A.B) + Ci.(A.B + A.B) \tag{3.19}
\]

\[
Co = A.B + Ci.(B + A)
\]

Since A and B tied together and the carry input is complementary signal of A and B, we assumed that \(A = B = X\) and \(Ci = \overline{X}\), then

\[
So = \overline{X}.(X + \overline{X}) = X \tag{3.20}
\]

\[
Co = X + (1 + \overline{X}) = X \tag{3.21}
\]

It shows that placing the complementary signals at the adder input stage results in complementary outputs to perform oscillation. Moreover, the adder based delay chain has to be initialized to complementary signals like \(A = B = X\) and \(Ci = \overline{X}\) respectively to perform oscillation. To accomplish this and to maintain the regularity of delay chain, the same adder circuitry is used to generate the complementary signals as shown below,

\[
\overline{Co} = A.B + Ci.(B + A) \tag{3.22}
\]
To initiate the adder oscillation, the external enable signal is used as shown below. When \(B = EN\) and \(A = 0\),

\[
C_o = C_i \cdot EN 
\]  

(3.23)

when \(B = \overline{EN}\) and \(A = 1\),

\[
\overline{C_o} = \overline{EN} + C_i 
\]  

(3.24)

The above expression represents the NAND and NOR functions, which is used to generate the complementary signals to initiate the delay chain. Thus, the adder based delay chain is initialized when the \(EN\) signal is turned \(OFF\) and it starts oscillating when it is turned \(ON\). It performs oscillation once its has been initialized to the steady state value as suggested by equation (3.23) and (3.24). Due to this, two adders placed at the input stage of the delay chain to output the complementary signals for the next stages to perform oscillation. Therefore, it really makes sense to analyze the FBB characteristics by using above explained adder based delay chain.
rather using simple inverter delay chain in the low-$V_{dd}$ region. Like the conventional ring oscillator, the 'n' number of 1-bit adder blocks are cascaded by rippling both the sum and carry signals to perform oscillation.

### 3.5 Effectiveness of Forward Body Bias

Short-channel effects and Drain-Induced Barrier Lowering are particularly severe in low-$V_{th}$ devices because of reducing the channel doping concentration to lower $V_{th}$, this causes the channel depletion depth to become larger. Further, the body-effect co-efficient $\gamma$ is lowered in short-channel transistor because the channel potential is influenced more by drain voltage than substrate due to the DIBL effect. Traditionally, the reverse-body bias (RBB) [Kes99] is used to reduce the leakage power, but the advantage of RBB is worsen with SCE and thus, the application of body biasing is shifted from RBB to FBB. Conventionally, forward body bias to the entire chip improves either performance or maintains the target frequency at reduced supply voltage especially in high-performance applications [Miy00]. Figure (3.10) represents the conventional method of body biasing in the CMOS logic gates using external body bias generators. FBB is applied to high-$V_{th}$ device to lower the $V_{th}$ down to the desired target value. Many recent research reports that the optimal FBB of 0.4-0.5V provides 13% and 37% performance improvement at 130nm and 90nm triple well process technology [Nar03] [Arn04] [Hua01].

Applying such forward body bias to the whole circuitry increases the performance at the cost of larger junction leakage capacitance and forward source-body junction current. Previous research work studied the FBB characteristics especially for high-performance microprocessor applications in the high-$V_{dd}$ region. However, in the preceding section we showed the effectiveness of FBB in the low-$V_{dd}$ region (the sub-threshold, near-threshold and the above-threshold region) to meet the target frequency requirements. It means that larger supply voltage reduction is possible for fixed performance targets as well as to mitigate the process fluctuations. In the following sections, the characteristics of various FBB configuration in the sub-threshold, near-threshold and above-threshold region using a 13-stage adder based delay chain are analyzed.
3 Energy Efficient Circuit Design

Figure 3.10: General architecture of body biasing in CMOS logic gates using external body bias circuitry

3.5.1 Selection of Optimal Forward Body Bias

Figure (3.11) illustrates the performance improvement offered by FBB at the elevated temperature in the low-$V_{dd}$ region. Increasing the FBB, lowers the threshold voltage and it improves the frequency of the delay chain till it reaches the optimum bias voltage, above which the frequency starts decreasing. Forward body bias greater than the optimum value rapidly increases the junction capacitance and drain-body junction forward current which limits further frequency improvement. Moreover, a threshold voltage reduction beyond the optimum value provides excessive diode forward current which reduces the full-swing output voltage and degrades the circuit noise margin. The figure (3.11) shows that the selection of optimum FBB in the low-$V_{dd}$ region depends on key factors like region of operation and the operating temperature. In the sub-threshold (0.2V) and above-threshold region ($\geq 0.5V$), the optimum FBB of 0.5V and 0.4V occurs at 25°C and 85°C respectively. However, the optimum FBB of 0.6V and 0.5V occurs at 25°C and 85°C in the near-threshold region (0.4V) respectively. These results indicate that, the same optimal FBB occurs when the adder based delay chain operates in the sub-threshold and above-threshold region at the elevated temperature. It further suggests that the desired optimum FBB of 0.4V and 0.5V attains the peak performance in all the three regions, when the delay chain operating at the maximum temperature of 85°C.

The positive temperature dependence in the low-$V_{dd}$ region improves circuit per-
3.5 Effectiveness of Forward Body Bias

Figure 3.11: The delay chain frequency increases with FBB in the sub-threshold, near-threshold and above-threshold regions at elevated temperature (All data points are compared to $V_{bs}=0V$ to find out how many times the frequency increase w.r.t ZBB)

Figure 3.12: Performance improvement of delay chain with optimum FBB at different operating voltage, all data points normalized to conventional Zero Body Bias scheme
formance, which is in contrast to conventional high-$V_{dd}$ scenario where the worst case delay occurs at higher temperature. Figure (3.12) shows the combined effect of positive temperature and forward body bias dependence in the low-$V_{dd}$ region to improve the frequency of the delay chain. Curve(I) and (II) show the frequency improvement offered by only optimal FBB compared to zero body bias at 25°C and 85°C respectively. And, Curve(III) represents the frequency improvement offered by the maximum temperature of 85°C compared to the room temperature without optimal FBB. It shows that optimal FBB has considerable performance improvement when it operates at room temperature compared to the maximum temperature of 85°C. However, the strong temperature dependency in the sub-threshold region without bias voltage improves the performance significantly compared to the optimal FBB cases operating at the elevated temperature. It shows that both optimal FBB and higher temperature operation in the low-$V_{dd}$ provides significant change in the performance compared to the conventional zero body bias scheme. By combining the advantages of both optimal FBB and the higher temperature operation together in the low-$V_{dd}$ region, significant performance could be achieved as shown in Fig. (3.12)(Curve(IV)). This section demonstrates that one could get significant performance benefits when both forward body bias and its temperature dependency combined together in the low-$V_{dd}$ region. Hence, the moderate amount of supply voltage reduction is possible for fixed performance targets, if the circuit under consideration adapts depending on the ambient temperature and operating region.

3.6 Transistor Sizing with Forward Body Bias

This section analyze the characteristics of FBB for minimum size and larger dimensioned 31-stage cascaded chain of identical inverters in the low-$V_{dd}$ region at the elevated temperature. In this section, we consider simple inverter based delay chain rather than adder based delay chain in order to have very regular structure to determine the effectiveness of FBB with transistor sizing. The optimal 0.4V FBB is applied to both pull-up and pull-down logic network with device width ratio of $\frac{W_p}{W_n} = 3$, the relative current driving capability varies and the corresponding intrinsic delay reduces due to smaller $ON$ resistance.
3.6 Transistor Sizing with Forward Body Bias

Figure 3.13: Performance improvement of 31-stage inverter based delay chain with optimum FBB and transistor sizing compared to conventional Zero Body Bias scheme

Figure (3.13) plots the intrinsic frequency and active power of 31-stage inverter delay chain ($FO = 1, C_L = 0$) as a function of device width, optimal FBB, temperature and supply voltage compared to the conventional zero body bias scheme. As the device size scales up from the minimum size device width ($W_n = 0.16\mu m$), the effective frequency improvement with optimal FBB is decreasing due to excessive source (or) drain to substrate diffusion capacitance. Moreover, the extra junction capacitance incurred by forward body bias exacerbates the total intrinsic capacitance, which actually lowers the efficiency of FBB for large dimensioned devices. On the other hand, the sub-threshold operation has lower capacitance compared to the near-threshold and above-threshold region since there is no gate-to-channel capacitances (when $V_{gs} < V_t$). At elevated temperature, it should be noted that the minimum size device in the sub-threshold region shows considerable frequency improvement over the larger-dimensioned devices because of lower diffusion and gate-to-channel
3 Energy Efficient Circuit Design

capacitances. In contrary, it shows comparatively minimal frequency enhancement in the near-threshold and above-threshold region. Therefore, it clearly suggests that the optimal FBB is highly efficient for minimum size devices when it operates in the sub-threshold region compared to the near-threshold and above-threshold region.

3.7 Noise Susceptibility with Forward Body Bias

In the ultra-low voltage regime, static noise immunity is becoming a metric of comparable importance to power and performance for the analysis and design of energy efficient circuits. This parameter allows us to determine the allowable noise voltage at the input of the gate so that the output logic state will not be corrupted [Han06]. Hence, it can be defined as the difference in magnitude between the faultless output and the input logic level of the victim circuit (circuit which is subjected to noise voltage from the external noise source). Thus, the noise immunity for HIGH logic state \( N_{IH} \) is given by \( N_{IH} = V_{OH} - V_{IH} \), where \( V_{OH} \) and \( V_{IH} \) represents the minimum HIGH output and input values. Similarly, for LOW logic state \( N_{IL} = V_{IL} - V_{OL} \), where \( V_{IL} \) and \( V_{OL} \) represents the maximum LOW input and output values.

To analyze the effect of these parameters, the minimum size cascaded inverter operated in the conventional zero body bias and forward body bias fashion in the elevated temperature is considered. Figure (3.14) represents the ratio of output LOW \( V_{OL} \) and HIGH \( V_{OH} \) logic level to the supply voltage for conventional zero body bias and optimal forward body bias at 25°C and 85°C respectively. It shows that the ratio of \( V_{OH} \) and \( V_{OL} \) increases exponentially below 0.2V \( V_{dd} \) irrespective of biasing voltage and operating temperature. It means that the logic circuits might lose its functionality when the static reverse biased diode leakage \( I_{junc} \) take over the switching sub-threshold leakage current \( I_{sub} \) as the supply voltage scales down below certain voltage.

At elevated temperature, the ratio of \( I_{sub}/I_{junc} \) gets worsen due to increased reverse-bias diode leakage which opposes the switching current to make full voltage swing in the conventional scheme. On the other hand, the FBB and its strong dependency on temperature exacerbates \( I_{sub} \) and \( I_{junc} \) in the ultra-low \( V_{dd} \) region.
3.7 Noise Susceptibility with Forward Body Bias

Figure 3.14: Simulated static noise margin $V_{OH}$ and $V_{OL}$ for a minimum size cascaded inverter with conventional zero body bias and forward body bias schemes.

Figure 3.15: Relative static noise margin of FBB and Conv. ZBB at various supply voltages.
Again, the increased forward bias source (or) drain junction leakage current \( I_{\text{junc}} \) fights with the sub-threshold switching current to maintain the full swing. This opposition between the switching and static current maintains the output LOW and HIGH voltage almost flat till 0.12V and below which the dominant \( I_{\text{junc}} \) lowers the output swing as shown in Fig. (3.15). Furthermore, the static LOW and HIGH noise immunity level is wide enough to tolerate the external noise voltage in the conventional scheme. However, it is observed that the optimum 0.5V FBB lowers the \( N_{IH} \) and increases the \( N_{IL} \) due to threshold voltage mismatch between the pull-up and pull-down devices. For the same 0.5V FBB, the larger pMOS threshold voltage reduction shifts the switching threshold towards the supply voltage which increases the input LOW \( V_{IL} \) and HIGH \( V_{IH} \) levels.

It shows that proper balancing between the nMOS and pMOS threshold voltage becomes necessary to attain better noise immunity level with forward body bias scheme. These results show that the output HIGH and LOW output voltage level becomes flat till certain supply voltage, and it will vary depending on the operating temperature and highly stacked circuits. Moreover, the circuit speed improved by FBB is suffered with lower noise immunity due to larger forward junction leakage in the low-\( V_{dd} \) region.

3.8 Forward Body Bias Configuration

3.8.1 Symmetric Forward Body Bias

In this section, the symmetric forward body bias is categorized into two main types, namely (1) Fixed FBB and (2) swapped FBB configurations. Figure (3.16) represents the conventional zero body bias, fixed and swapped FBB schemes with symmetric body bias respectively. In fixed FBB scheme, the constant optimal FBB is applied to both pull-up and pull-down logic network symmetrically irrespective of operating region using triple-well process technology. This optimum FBB is generated by an external body bias circuitry [Miy00], it supplies optimum FBB and zero body bias during active mode and standby mode of operation respectively as shown in Fig. (3.17). This system is very simple because it does not require any switch cells and negative-voltage generator. Conventionally, such systems are employed in
high performance microprocessor applications with additional circuitry to adapt the circuit performance depending on process parameters. Such method guarantees the optimal trade-offs between the extra layout area for body-to-source signal generation and the performance advantages. This approach exacerbates the unwanted active leakage power from the forward body-biased non-evaluation devices, when optimum FBB is applied to the entire state-of-the-art static CMOS logic circuits. This extra active-leakage power and area overhead is negligible for high performance targets, but it mounts significant burden for low energy applications.

To overcome this area and power overhead imposed by external body bias generators, an internal forward body bias approach using supply voltage as the body bias voltage is proposed for ultra-low voltage circuits like B-DTMOS and LVSB approaches [Shi98] [Nar04]. This alleviates most of the limitations of fixed FBB at the expense of performance loss due to lower supply voltage. In this swapped FBB scheme, the body-to-source voltage is varied in accordance to the supply voltage ($V_{bs} = V_{dd}$) to constantly improve the performance which is depicted in Fig. 3.16(c). It means that, the ratio of change in threshold voltage to the corresponding change in FBB varies with the region of operation. Since the FBB is equal to supply volt-
Figure 3.17: Schematic cross-section of Forward Body Bias system to generate $V_{bsn}$ and $V_{bsp}$ respectively (Mas05)

age, the ($I_{on}/I_{off}$) of swapped FBB is smaller than that of fixed FBB due to smaller body bias voltage when it operates below the optimal FBB value. However, this approach is limited to maximum supply voltage of 0.6V (diode cut-in), to avoid the parasitic drain/source junction diode to turn on heavily. It shows that, the swapped FBB is most likely suitable for ultra-low voltage applications than the conventional fixed FBB schemes, because of smaller area overhead and less design complexity.

### 3.8.2 Asymmetric Forward Body Bias

Applying symmetric FBB to the entire static CMOS circuits increases the active leakage power from non-evaluation transistors. For example, if we consider a simple 2 input NAND gate with input vector "11" or "00", only half of the logic devices make the evaluation transition and the other half contributes to static leakage current. It means that applying FBB to the entire circuit creates unwanted active leakage power from the non-evaluation transistors [Jay07]. On the other hand, it is also impractical to place low-threshold and high-threshold transistors depending on the evaluation and non-evaluation transitions using multi-threshold voltage CMOS technology. This shows the inadequacy of using symmetric FBB and dual-threshold voltage technique in state-of-the art static CMOS circuits to obtain low-leakage
3.8 Forward Body Bias Configuration

power, which will be discussed in detail in the next chapter.

Figure 3.18: (a) Asymmetric FBB with well-contact tied together (b) Asymmetric n-well Forward Bias - only pull-up devices are forward biased \((V_{AFB}=0\,V)\) (c) Asymmetric p-well Forward Bias - only pull-down devices are forward biased \((V_{AFB}=V_{dd})\)

To overcome these difficulties, the self-adaptive forward body bias depending on the output node voltage is proposed, in which devices are biased depending on the evaluation transition [Jay06]. In this scheme, the forward body bias \((V_{bs} = V_{dd})\) is applied to either pull-up or pull-down logic network depending on the output evaluation transition. A few guidelines arising from these prior studies such as (1) application of FBB depending on the evaluation transition improves performance with low active leakage power in the low \(V_{dd}\) regime and (2) the static dual-threshold voltage technique may not be a suitable choice for ultra-low voltage applications due to lower static noise immunity. From these guidelines and minimum energy-per-cycle perspective, we analyze the asymmetric forward body bias concept in which the high threshold pull-up and pull-down devices are forward body biased asymmetrically by connecting the well contacts as shown in Fig. 3.18(b&c).

Conventionally, the pMOS in n-well and common p-substrate of nMOS devices are connected to \(V_{dd}\) and \(gnd\) rails respectively as shown in Fig.3.16(a) to keep the parasitic diodes in reverse-bias mode. In this approach, the well contact is connected either to the substrate potential \((V_{AFB} = 0\,V)\) called as n-well forward
bias or to the supply voltage \( V_{AFB} = V_{dd} \) called as \textit{p-well forward bias} respectively as shown in Fig. 3.18(b&c). This arrangement lowers the threshold voltage of either pull-up or pull-down logic network and correspondingly increases the low-to-high or high-to-low evaluation transition rate respectively. Further, the respective high-threshold voltage of pull-down and pull-up logic network maintains the static current under control. It means that lowering the \( V_{th} \) in one direction improves the driving strength and maintaining the high-\( V_{th} \) provides lower static current in the other direction. Thus, this approach improves performance by making one of the transition faster than the other with lower static leakage current in the ultra-low voltage regime. Similarly, [Mog92] proposed a new CMOS approach in which common local well contact is enabled by single substrate supply voltage and [Bry01] proposed CMOS inverter by matching the nMOS and pMOS off-currents by biasing the well potential by a simple regulator output. Hence, in our case we studied the asymmetric FBB configuration to figure out the optimum between the energy-per-cycle and performance in the low-\( V_{dd} \) regime which will be discussed in the following sections.

3.8.3 Power and Performance Comparison

(a) Symmetric Forward Body Bias

Figure (3.19) illustrates the performance improvement and the corresponding total power consumption of minimum size 13-stage adder based delay chain for fixed, swapped FBB and the conventional zero body bias schemes at the elevated temperature. The average active power and static power of the delay chain is evaluated when the enable signal \( EN=1 \) (delay chain performs switching) and \( EN=0 \) (delay chain reached the steady state) respectively. As can be seen, both the performance and power of the delay chain increases with applied forward body bias.

The fixed FBB scheme exacerbates both the power and performance of the delay chain with 0.4V optimal FBB. When the delay chain operates below 0.5V \( V_{dd} \), significant threshold voltage reduction enhances performance at the cost of larger static current due to extensive source/drain to body diode leakage current as discussed in section (3.3.2). In the case of the swapped FBB scheme, the moderate \( V_{th} \) reduction due to the smaller body bias improves the performance considerably with
minimal increase in power consumption. For example, in the sub-threshold region ($V_{dd} = 0.2V$), the fixed and swapped FBB scheme improves performance by 2.6X and 1.8X at the expense of 3.2X and 2.2X extra power compared to conventional zero body bias respectively (Here, 'X' represent "times"). Similarly, both schemes improve performance by 1.3X and the corresponding 1.5X extra power increase in the above-threshold region ($V_{dd} = 0.5V$). At elevated temperature, the static power is dominant over active power due to its exponential dependency on diode leakage current. This significantly increase the overall power consumption of fixed FBB scheme compared to the swapped FBB and zero body bias scheme in the sub-threshold region. Moreover, the performance improvement at higher temperature with forward body bias is relatively low compared to the room temperature. Since the body
bias of swapped FBB scheme $V_{bs}=V_{dd}$ which keeps the parasitic drain-to-substrate diode at zero bias and further it minimally forward biases the source-to-substrate diode due to smaller body bias. Hence, the swapped FBB scheme has lower power consumption with moderate performance improvement compared to the fixed FBB scheme.

Moreover, for fixed performance targets one could attain considerable supply voltage reduction when operating in the low-$V_{dd}$ regime. For iso-performance operation, fixed FBB scheme with 0.4V optimal value reduces $V_{dd}$ approximately by 42mV compared to conventional zero body bias scheme. On the other hand, the swapped FBB scheme shows around 25mV $V_{dd}$ reduction. This means instead of operating at 200mV with conventional ZBB one could set the $V_{dd}$ of 175mV, since $V_{bs}=V_{dd}$, this reduction will vary depending on the operating voltage. Thus, on average around 12% and 9.5% supply voltage reduction is obtained for iso-performance operation with fixed and swapped FBB respectively compared to the conventional scheme. One can expect that, the total power consumption should reduce with supply voltage, however the parasitic diode leakage current and respective higher junction capacitance minimizes the advantage of reducing the supply voltage for active power reduction as shown in Fig. (3.19). On the other hand, one can get considerable total power reduction by withdrawing the FBB in the standby mode of operation. For example, instead of operating at 300mV if one operates at 258mV with 0.4V FBB the static power at this potential is lower than that of 300mV operation. Moreover, there is no advantage if we go down below 200mV operation because static power becomes dominant over active power consumption.

(b) Asymmetric Forward Body Bias

Figure (3.20) illustrates the performance improvement of minimum size 13-stage adder based delay chain with asymmetric FBB compared to zero body bias at the elevated temperature. It shows that substantial threshold reduction due to n-well forward bias exacerbates the frequency of the delay chain compared to the p-well forward bias. It means that forward biasing only the PUN and zero biasing the PDN escalates the frequency compared to forward biasing only the PDN and zero biasing PUN scenario. Like symmetrical FBB, there exists an optimum asymmetric
3.8 Forward Body Bias Configuration

FBB at which the frequency of the delay chain reaches a maximum. This opti-

Figure 3.20: The 13-stage adder delay chain frequency is improved with asymmetric FBB compared to the conventional ZBB scheme at the elevated temperature

mum value depends on several parameters like transistor width, number of fan-in, circuit architecture and the operating temperature. For instance, at 0.3V $V_{dd}$ the peak performance of the delay chain occurs when the PDN and PUN are forward body biased unequally with 0.1V and 0.2V respectively. It indicates that the peak performance point is reached when the entire circuitry is forward biased asymmetrically. Thus, it reveals that optimum 0.1V AFBB is chosen depending on the peak performance in the sub-threshold, near-threshold and above-threshold region at the elevated temperature. However, this optimum value is not uniform between PUN and PDN and this might incur larger threshold voltage fluctuation in the low-$V_{dd}$ regime. To overcome this, we chose $V_{AFB} = V_{dd}/2$ as the second optimal value which provides symmetrical FBB to both PUN and PDN which is highly important for low-voltage operation. This will be discussed in detail in the next chapter.
Like symmetric FBB, the efficiency of AFBB decreases at higher temperature in improving the performance of the delay chain compared to the room temperature. Moreover at 25 °C, both the n-well and p-well forward bias shows considerable performance gain compared to zero body bias scheme. Whereas at high temperature 85 °C, the frequency improved by p-well forward bias is reduced even lower than the conventional zero body bias scheme. This clearly suggests that both n-well forward bias and chosen 0.1V AFBB is highly efficient in improving the performance compared to p-well forward bias and zero body bias when it operates at the elevated temperature.

**Figure 3.21:** The change in power consumption with asymmetric FBB compared to conventional ZBB scheme at the elevated temperature.

Figure (3.21) represents the change in average active and static power with AFBB compared to conventional zero body bias at the elevated temperature. The simulation result indicates that both active and static power has increased substantially with AFBB due to excessive source/drain to body diode junction leakage current.
3.8 Forward Body Bias Configuration

compared to conventional zero body bias scheme. However, this increased active and static power descends with AFBB when the PUN switches from FBB to ZBB mode and PDN switches from ZBB to FBB mode respectively at the cost of performance. Further increase in AFBB, reduces the active switching power because of reverse body bias nature of pull-up pMOS devices. Even though the threshold voltage of the pull-up devices are increased, the excessive forward biased PDN enlarges the static power when \( V_{AFB} \geq V_{dd} \). Moreover, the static power of the n-well forward bias is higher than p-well forward bias because the low-threshold pull-up device draws a larger static current compared to its counterpart. On the other hand, the zero body biased nature of pull-up logic resists the static current flow and thus the p-well forward bias has lower static power than its counterpart. From these discussion, one can expect that there should be an minimum active and static power with AFBB when it operates in between the n-well and p-well forward bias schemes. Thus, this optimum is chosen as 0.1V AFBB and \( V_{AFB} = V_{dd}/2 \) when it operates in asymmetrical and symmetrical fashion respectively. If we look at each design corners, the \( V_{AFB} = V_{dd}/2 \) provides the optimal solution compared to 0.1V AFBB in the low-\( V_{dd} \) regime.

3.8.4 Minimum Energy-Per-Cycle Comparison
(a). Symmetric Forward Body Bias

The energy-per-cycle is estimated by multiplying the total power consumption with the delay of the minimum-size adder chain as shown in Equation (3.9). It can be defined as, the amount of energy required to perform one complete cycle (50% duty cycle). It combines both active energy and static energy-per-cycle, which gives the amount of energy spent during active switching and static non-switching operation. Figures (3.22) and (3.23) show the energy-per-cycle for conventional zero body bias, symmetric and asymmetric FBB at 25°C and 85°C respectively. It shows that the energy-per-cycle exacerbates with FBB due to increased source/drain to body junction leakage current. Lowering threshold voltage due to FBB reduces the circuit delay at the expense of increasing the static leakage power incurred by FBB. In addition, the strong temperature dependency of FBB below 0.5V \( V_{dd} \) further increases the overall energy-per-cycle.
In case of symmetric FBB, the minimum energy-per-cycle occurs at 0.2V and 0.3V $V_{dd}$ for conventional zero body bias and swapped FBB scheme in the sub-threshold region while it operates at 25°C and 85°C respectively. In the case of fixed FBB, the optimal 0.4V FBB increases the minimum energy-per-cycle to 0.3V and 0.4V $V_{dd}$ at 25°C and 85°C respectively. It shows that the applied FBB and the operating temperature determines the minimum energy-per-cycle, the supply voltage corresponding to $E_{min}$ called as minimum operating supply voltage ($V_{ddmin}$) as shown in equation (3.14). The static leakage energy becomes dominant when the supply voltage scales further below $V_{ddmin}$.

(b). Asymmetric Forward Body Bias

Like the conventional ZBB scheme, the minimum energy/cycle occurs at 0.2V and 0.3V for n-well, p-well and the optimal 0.1V AFBB when it operates at 25°C and 85°C respectively, which is shown in Fig. (3.23). This energy increase due to optimal 0.1V AFBB is minimal compared to the n-well and p-well forward bias at 0.2V $V_{ddmin}$. This is exactly the point where both PUN and PDN are forward biased symmetrically when 0.1V is applied to the common well-contact. This actually
proves that the optimal value of $V_{AFB} = V_{dd}/2$ is more beneficial than operating at 0.1V AFBB optimal value. It clearly indicates that conventional sub-threshold circuit operation without any body bias provides the minimum energy-per-cycle, but at the cost of severe performance penalty. However, the applying body bias either using external or internal body bias generator improves performance in the low-$V_{dd}$ regime at the expense of energy increase. Hence, it shows there exist a clear trade-off between the energy-per-cycle and performance metric when sub-threshold operation and FBB combined together in the ultra-low $V_{dd}$ region. At 0.2V $V_{dd_{\min}}$, the fixed and swapped FBB improves performance by 2.6X and 1.8X at the expense of 1.6X and 1.2X energy increase compared to the conventional scheme as shown in Fig. (3.24). Further, there is no advantage in energy reduction for fixed performance target (iso-performance) below $V_{dd_{\min}}$ due to dominant static energy. On the other hand, the performance improvement and the respective energy increase is minimal in asymmetric FBB compared to symmetric FBB schemes. At the elevated temperature, the symmetric FBB increases the minimum supply voltage with significant performance increase and hence it is not a viable method for low-energy applications. From these, we can conclude that symmetric FBB with $V_{bs} = V_{dd}$ and $V_{bs} =
$V_{dd}/2$ provides considerable performance gain with minimal energy/cycle increase in the ultra-low $V_{dd}$ regime. Further, one can think of possible active and static energy reduction by lowering the supply voltage for iso-performance solution during active mode and withdrawing the FBB during standby mode of operation.

Figure 3.24: Energy-per-cycle and the corresponding performance improvement of 13-stage adder based delay chain for symmetric and asymmetric FBB schemes at the elevated temperature.

Generally, the energy-per-cycle metric targets mainly ultra-low energy and ultra-low performance solutions. However, to determine the energy efficient design we need a metric that has equal priority for both power and delay terms. Therefore, we considered the energy-delay-product (EDP) metric which is the product of equation(3.13) and equation(3.10), i.e., $EDP = power\text{-}delay\text{-}product \times delay$[Cha92]. To improve this metric, either the performance has to be increased or the energy should be reduced without affecting the other parameters in the EDP metric. As we know, the applied FBB lowers threshold voltage which reduces the circuit delay, on the
other hand it increases static energy due to junction leakage current. Hence, to analyze this we study the energy-delay-product metric with various forms of FBB schemes at elevated temperature in Fig. (3.25). As expected, at 25°C the energy-delay-product is improved for both symmetric and asymmetric FBB schemes due to the larger circuit delay reduction by FBB. However, at higher temperature (85°C), the above-mentioned schemes increase EDP because of stronger dependency of FBB on the junction leakage current and temperature. Though it increases performance at elevated temperature, the static energy exacerbates with higher temperature and applied FBB, which increases the overall energy-delay-product. As we discussed in the previous section, at minimum supply voltage the fixed FBB improves the frequency of the delay chain by 2.6X with 1.6X reduction in EDP compared to conventional scheme. Since FBB and its dependency on temperature reduces the circuit delay, the EDP metric always improves with the applied FBB at room temperature. At 25°C, both symmetric and asymmetric FBB schemes reduce the EDP with increased frequency in the ultra-low $V_{dd}$ regime.

### 3.9 Analysis of Process Variations

Process-induced MOSFET parameter variations [Pel89] play a vital role in ultra-low voltage, low-power applications. Especially, it has higher sensitivity towards sub-threshold operation where minimum energy is the primary criteria than the performance metric. This is mainly due to the exponential dependency of the sub-threshold current on $V_{th}$ as given in Equation (3.11). As we discussed in the previous section, the minimum supply voltage ($V_{dd_{min}}$) and the respective minimum energy/cycle ($E_{min}$) changes with temperature and the given FBB. With $V_{dd_{min}}$ constant at 0.2V, the corresponding $E_{min}$ increases with the FBB at 25°C. However, this analysis has been done only for temperature variation and it did not take other process parameter variations into consideration. Since the sub-threshold current has exponential dependency on $V_{th}$, even a minor change will significantly affect its energy efficiency. [Zha05] observed a variation in gate delay as high as 300% from nominal value, when operating in the sub-threshold region. Hence, it becomes difficult to meet specifications without designing with a significant margin.
which minimizes the efficiency of low power systems. Hence, in the following section we discuss the influence of process variations especially in the sub-threshold region (since a $V_{dd_{min}}$ and $E_{min}$ exist when $V_{dd} < V_{th}$) and the possible approaches to mitigate these variations.

Due to the exponential dependency of $V_{th}$ and $L_{eff}$ on sub-threshold current, the influence of process variation exacerbates in sub-threshold region. An substantial research work has been done in the past to study the sources of variation and proposed several techniques for design for variability in the presence of inter-die and intra-die variation for high-performance microprocessors [Nas00, Tsc02, Bor03, Eis97]. Most of these methods mitigate only with-in die variation by adapting the body bias by tracking the performance level of the chip, in which both FBB and RBB fed into
slow and fast running parts of the chip, respectively. These variations have been studied further to sub-threshold operation with transistor-sizing, introducing larger logic depth and optimal pipelining strategy [Zha05]. This study showed considerable energy reduction for intra-die process variation by introducing larger logic depth between the pipelining stages. In order to extend our study further, in the following section we present the worst-case analysis supported by manufacturing process characterization and Monte-carlo analysis method to determine the statistical spread of desired parameter of interest using body bias and transistor sizing methodologies.

3.9.1 Sensitivity Analysis with Threshold voltage

Sub-threshold circuits are highly prone to process variation because of their exponential dependence on threshold voltage and the effective channel-length dependency on sub-threshold current. The transistor threshold voltage is mainly determined by device parameters set during manufacturing process and it is heavily affected by imperfections in the process steps. Thus, the standard deviation of the manufactured threshold voltage from its mean design value can be expressed as [Sto98]

$$\sigma_{V_{th}} = \frac{t_{ox}\sqrt{NT}}{\sqrt{W_{eff}L_{eff}}}$$

where $t_{ox}$ is the thickness of the gate oxide, $N$ is the channel doping density, $T$ is the absolute temperature and $W_{eff}$ and $L_{eff}$ are the effective width and length of the transistor, respectively. This deviation causes variation in $V_{th}$ which contributes to variability in gate delay and energy-per-cycle, as we have seen before in Equation (3.10) and (3.13) respectively. From a designer’s perspective, these delay and energy-per-cycle spread is varied by many distinct sources: fluctuations in power supply voltage, temperature, noise coupling among nets due to environmental conditions and the processing imperfections in physical device structure ($L_{eff}$, $t_{ox}$, $W$ and $V_{th}$). Thus, the delay spread increased relative to its nominal value because of these distinct parametric variation.

3.9.2 Worst-case Corner Analysis

For worst-case design analysis, we simulated a 13-stage adder based delay chain for all different process corners like Slow-Slow (SS), Fast-Fast (FF) and typical
Figure 3.26: Worst-case analysis of power and frequency with different process corners using 13-stage adder based delay chain

(TT) to analyze the performance spread of our design under more realistic process variations. The corner analysis is a worst-case approach, where we can simulate over multiple corners of process, power supply, and temperature to get an insight of system behavior before being manufactured [Bolt91]. Figure (3.26) shows the corner analysis of 13-stage adder delay chain with different process corners in terms of power and frequency of the delay chain. It shows that there is clear minimum performance and power in the nominal case, but there is no minimum in the worst-case power (FF) and performance (SS) corners. Higher temperature and its positive dependency below zero-temperature coefficient exacerbates the frequency and power of the delay chain, and hence it is not suitable for ultra-low power applications. As we know, the worst case (SS) corner provides ultra-low performance and the correspondingly
ultra-low power due to higher $V_{th}$ from the nominal $V_{th}$. To restore the $V_{th}$ to its nominal value, the optimal 0.4V FBB is fed into the delay chain which brings the $V_{th}$ to near-nominal value to meet the design specifications. The increased power consumption due to higher junction leakage current is still minimal when FBB is applied to worst-case performance corner.

Figure 3.27: Worst-case analysis of 13-stage adder based delay chain in terms of energy-per-cycle, supply voltage and frequency of the delay chain

Figure (3.27) shows the variation in energy-per-cycle with respect to different process corners for the same 13-stage adder delay chain. As one can see that the minimum energy-per-cycle ($E_{min}$) and the corresponding minimum supply voltage ($V_{ddmin}$) changes rapidly with worst-case corners, operating temperature and the applied FBB potential. It shows that, this traditional worst-case analysis does not
provide good insight about the circuit fluctuations in the ultra-low voltage regime (esp. in the sub-threshold operation) because it only consists of FAST and SLOW parameter-sets which are used to simulate maximum and minimum performance levels. In this, all design parameters are set to maximum and minimum limits which results in maximum (FF) and minimum (SS) circuit power and performances. Such correlated parameter-sets results in circuits being over-designed which unnecessarily exacerbates the energy-per-cycle metric.

Hence, the worst-case corner analysis is not an effective way of estimating process parameter fluctuations for sub-threshold operation. Therefore, in the following section we discuss the statistical variation using Monte-carlo simulation methodology to determine the delay and energy variability in the ultra-low voltage regime.

3.9.3 Study of Random Dopant Fluctuations (RDF)

As the operating supply voltage is brought below $V_{th}$, the threshold voltage variation due to Drain-Induced-Barrier-Lowering (DIBL) is minimal compared to higher drain-to-source voltages. As equation (3.25) suggests the effective channel length variation $L_{eff}$ due to DIBL and Short-channel effects (SCE) is relatively smaller in the sub-threshold operation due to the reduced supply voltage compared to the above-threshold region. However, as the device dimension scales down into the nanometer regime, the $V_{th}$ variation caused by random dopant fluctuation [Ase01] is dominant over other process parameter variations for sub-threshold operation. Since $I_{sub}$ has exponential dependency on $V_{th}$ and inversely proportional to $L_{eff}$, the $V_{th}$ variation has significant impact due to random doping fluctuation (RDF) and the channel length reduction due to technology scaling. Thus, the threshold voltage variation depends on the random variation of number and position of dopant atoms as the effective channel length scales down as evident in Equation (3.25). Therefore, the sub-threshold current with respect to the threshold variation can be expressed as follows using equation (3.11):

$$I_{sub} \propto e^{(V_{gs}-V_{th})/m\sqrt{T}}$$

(3.26)
Differentiating the above equation with respect to $V_{th}$, and then dividing by $I_{sub}$ on both side yields,

$$\frac{\partial I_{sub}}{I_{sub}} \propto \frac{(V_{gs} - V_{th})}{mV_T} (-1) \propto \frac{V_{th}}{mV_T}$$  \hspace{1cm} (3.27)

This equation shows how the sub-threshold current varies from its nominal value in the presence of threshold voltage variation. As we can see, the $I_{sub}$ variation is directly proportional to the $V_{th}$ variation and inversely proportional to the sub-threshold swing and thermal voltage. This uncertainty exacerbates since the sub-threshold swing factor ‘m’ reduces when the supply voltage is reduced below $V_{th}$. Thus, the sub-threshold current spreads from its nominal value because of $V_{th}$ variation introduced by RDF and decreasing sub-threshold swing factor in the sub-threshold operation.

From this, one can determine the delay and the corresponding energy variability in the sub-threshold region. As $V_{dd}$ scales below $V_{th}$, the change in the exponential dependency of circuit delay with respect to threshold voltage can be written by combining equation (3.10) and (3.26) together as follows

$$\frac{\partial t_d}{\partial V_{th}} \propto \alpha C_s V_{dd} \left( \frac{V_{dd} - V_{th}}{mV_T} \right) \frac{1}{e^{\frac{(V_{dd} - V_{th})}{mV_T}}}$$  \hspace{1cm} (3.28)

Dividing the above equation by $t_d$ on both sides, yields

$$\frac{\partial t_d}{t_d} \propto \left( \frac{V_{dd} - V_{th}}{mV_T} \right)$$  \hspace{1cm} (3.29)

Equation (3.28) and (3.29) show the circuit delay variation with respect to the threshold voltage and the respective variation from the nominal delay value. Like equation (3.27), the delay uncertainty is proportional to the threshold voltage and the sub-threshold swing factor in the sub-threshold region. Thus, it shows that the delay uncertainty introduced by the threshold voltage variation causes the respective energy-per-cycle to vary as evident in Equation (3.12), this is discussed in detail with simulation results in the following sections.

### 3.9.4 Simulation Results and Discussion

So far, we addressed the process fluctuation in the sub-threshold region with transistor sizing and forward body bias to minimize the effects of random dopant fluctuations (RDF) and short-channel effects. Increasing device width reduces RDF,
because the sub-threshold current increase with larger device width minimizes the effect caused by random doping variations. On the other hand, the source/drain to body depletion width reduces with FBB, which decreases the short-channel effects due to technology scaling. Figure (3.28) shows the $3\sigma/\mu$ delay and active energy variability of 31-stage inverter based delay chain versus device width, forward body bias with supply voltage using Monte Carlo simulation methodology. Each inverter is equally loaded ($F_O=1$) and it is enabled by an external NAND gate to perform oscillation. As device width increases, the relative delay variability and the respective energy variability decreases compared to the minimum size device, as expected.

![Figure 3.28: ($3\sigma/\mu$) 31-stage inverter based delay chain showing the delay and active energy variation with optimal FBB and various transistor dimensions](image)

At minimum supply voltage $V_{dd_{min}}$ of 0.2V, the minimum-size device shows nearly 35% and the corresponding energy variability of 11% respectively, under no body bias condition. By increasing the device size to 10X of the minimum size device, the delay and energy variability drops to 25% and 7% respectively without any body bias voltage. If we combine both the optimal forward body bias and device sizing
3.9 Analysis of Process Variations

Table 3.3: Mean ($\mu$), standard deviation ($\sigma$), delay and power variability ($3\sigma/\mu$) for conventional zero body bias, asymmetric FBB and symmetric FBB schemes respectively in the above-threshold voltage region ($0.5V_{dd}$)

<table>
<thead>
<tr>
<th>Body-Bias schemes</th>
<th>Delay</th>
<th>Active Power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu$ (ns)</td>
<td>$\sigma$ (ns)</td>
</tr>
<tr>
<td>Conv.ZBB</td>
<td>3.59</td>
<td>0.08</td>
</tr>
<tr>
<td>Asymmetric-FBB</td>
<td>3.77</td>
<td>0.33</td>
</tr>
<tr>
<td>Symmetric-FBB</td>
<td>3.62</td>
<td>0.05</td>
</tr>
<tr>
<td>$V_{bs}=V_{dd}$</td>
<td>3.64</td>
<td>0.05</td>
</tr>
<tr>
<td>Symmetric-FBB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{bs}=0.4V$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Together, the maximum delay and energy variability reduction of 23% and 6% is achieved. As we discussed earlier, the delay variability introduced by process fluctuation in the sub-threshold region is larger than that of above-threshold region. At 0.5V supply voltage, the delay and energy variability drops from 15% to 10% and 2% to 1% with both optimal FBB and larger device size together in the above-threshold regime. This $3(\sigma/\mu)$ variability reduction is achieved at the expense of higher active energy due to larger device size and junction leakage current incurred during body bias. This shows that by combining both device sizing and FBB together, one can get considerable variability reduction at the expense of energy increase which is inevitable.

Table 3.4: Mean ($\mu$), standard deviation ($\sigma$), delay and power variability ($3\sigma/\mu$) for conventional zero body bias, asymmetric FBB and symmetric FBB schemes respectively in the sub-threshold voltage region ($0.2V_{dd}$)

<table>
<thead>
<tr>
<th>Body-Bias schemes</th>
<th>Delay</th>
<th>Active Power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\mu$ (ns)</td>
<td>$\sigma$ (ns)</td>
</tr>
<tr>
<td>Conv.ZBB</td>
<td>349.03</td>
<td>198.98</td>
</tr>
<tr>
<td>Asymmetric-FBB</td>
<td>43.59</td>
<td>160.7</td>
</tr>
<tr>
<td>Symmetric-FBB</td>
<td>98.97</td>
<td>122.52</td>
</tr>
<tr>
<td>$V_{bs}=V_{dd}$</td>
<td>263.72</td>
<td>73.85</td>
</tr>
<tr>
<td>Symmetric-FBB</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$V_{bs}=0.4V$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
To extend our study further for different body bias configuration in the presence of intra-die process variation, we again consider the minimum size 13-stage adder based delay chain as shown in Fig. (3.8) using Monte-Carlo simulation environment. The simulation has been done using randomly generated samples (N=50) with $W_p = 3W_n$ device matching ratio. Table (3.3) and (3.4) shows the Mean ($\mu$), standard deviation ($\sigma$), delay and power variability ($\frac{3\sigma}{\mu}$) for conventional zero body bias, symmetric and asymmetric FBB schemes in the above-threshold and sub-threshold voltage region respectively. As we expect, the ($\frac{3\sigma}{\mu}$) delay and active power variability is lower in above-threshold region compared to the sub-threshold region. The asymmetric FBB scheme with the chosen optimal 0.1V shows significant delay variation compared to other schemes, this is mainly due to the $I_{on}$ mismatch between nMOS and pMOS devices introduced by the unequal bias voltage. Therefore, the optimal bias voltage of $V_{dd}/2$ is considered because it provides equal $I_{on}$ which minimizes the delay ![](https://example.com/delay.png)

Figure 3.29: ($3\sigma/\mu$) 13-stage adder delay chain showing the delay and active energy variation at sub-threshold and above-threshold region for various body bias configuration
3.9 Analysis of Process Variations

variation significantly compared to asymmetric optimal FBB of 0.1V as discussed in Section (3.8.3). Thus, choosing the symmetric forward body bias would be more beneficial than the asymmetric FBB for low-voltage circuit operation. Further, the amount of body bias applied determines the optimum between the energy increase w.r.t FBB and the corresponding variability reduction.

Figure (3.29) shows the delay and active energy variability for 13-stage adder delay chain at sub-threshold and above-threshold region for various body bias configuration at 25°C. It shows that, the variability due to asymmetric FBB of 0.1V is significant in both sub-threshold and above-threshold regions. As we discussed before, the optimal 0.4V FBB shows significant delay and active energy variability of 51% and 66% compared to the conventional zero body bias scheme in the sub-threshold voltage region. At above-threshold region (0.5V), the variability reduction of 15% and 10% in delay and active energy is obtained with the symmetric optimal FBB. It clearly shows that optimal FBB reduces the delay variability and the respective active energy compared to other body bias schemes in the ultra-low voltage regime. However, the optimal 0.4V FBB increases the junction leakage current, but this increase is minimal compared to the variability reduction.

In this section, we discussed the characteristics of forward body bias in the ultra-low voltage regime (esp. $V_{dd} \leq 0.5V$). Different forward body bias schemes like symmetric and asymmetric configurations are presented in terms of energy-per-cycle and performance metric. In addition, the worst case design analysis and process parameter fluctuation using Monte-Carlo simulation are discussed to mitigate process fluctuation in the sub-threshold operation.
3 Energy Efficient Circuit Design
4 Low-Voltage Fine-grained FBB

For ultra-low-energy processors, the conventional method of reducing supply voltage improves energy efficiency is limited by severe delay penalty and dominant leakage power in the ultra-low voltage regime. In the nanometer era, the static sub-threshold and gate-leakage currents are predicted to be significant and are expected to have a substantial amount in the total power consumption. Hence, to analyze the energy-per-cycle with performance improvements for energy efficient design, such circuit optimization techniques which combine the monotonic precharge-evaluate logic circuits and application of forward body bias depending on the evaluation and idle modes within logic stages become interesting. Therefore, in this section we propose and evaluate different intrinsic forward body bias configurations for wide range of precharge evaluate logic circuits to trade-off between energy and performance in the ultra-low voltage regime.

4.1 Introduction

Energy efficient LSIs are primary concern in today’s battery enabled ubiquitous computing portable devices. As discussed in the previous section, applications like autonomous wireless sensor nodes [Jay05] in which the energy efficiency is the main focus for extending the battery life time will opt for sub-threshold operation at minimum energy consumption with severe performance penalties. Improving energy efficiency in the traditional way of lowering the supply voltage is limited by the large portion of leakage power in the total power consumption, severe delay degradation and circuit failure at very low $V_{dd}$ [Nar04]. However, the performance targets and proper circuit operation at lower supply voltage are obtained by applying forward body bias using expensive triple-well process technology. In this conventional forward body bias approach, an extra body bias circuitry is required to generate the
constant body-to-source voltage to achieve the required target performance. Moreover, careful design should be done to keep the body-to-source signal generation and routing power lower than the performance advantages. To prevent this area overhead and to improve energy efficiency, Narendra et al. proposed low-voltage swapped-body biasing technique in alleviating the limits to supply voltage reduction and energy improvements achievable by conventional schemes. However, this scheme has higher leakage power during active mode of operation because of low-threshold voltage non-evaluation devices. In addition, reversing the supply rails to zero body bias mode during the standby mode of operation is also not practically feasible. This approach improves the energy efficiency by lowering delay during active mode but at the expense of both the active and standby leakage power which may not be suitable for energy efficient design.

Furthermore, the main limitation of applying forward body bias globally to traditional static CMOS circuits is unwanted active leakage power, which arises from low-threshold transistors during non-evaluation circuit stage. To overcome these issues, the selective FBB depending on the evaluation transition is also impractical in the conventional static CMOS circuits. Therefore, a selective fine-grained forward body bias called pre-charge node based asymmetric forward-body-bias (PNFBB) technique is proposed in which bias voltage is applied depending on the evaluation transition for precharge-evaluate logic circuits.

Section (4.2) discusses the asymmetric forward body bias scheme, section (4.3) discusses the precharge-node based asymmetric forward body bias technique, section (4.4) discusses static and dynamic precharge-evaluate logic circuits, section (4.5) discusses the synthesis of precharge-evaluate logic circuits and section (V) evaluates the simple AND-OR simulation results of the proposed approach with previous techniques using industrial 130nm process technology.

4.2 Asymmetric Forward Body Bias - An overview

Conventionally, the method of applying forward body bias symmetrically to the entire circuitry improves performance at the expense of larger junction leakage current and extra routing area overhead. To analyze these parameters, we evaluate
the energy efficiency by applying forward body bias asymmetrically to the entire static CMOS by connecting the n-well and p-well contacts together to make one of the transition faster than the other. [Bry01] proposed CMOS inverter to match the nMOS and pMOS off currents by connecting the well-contacts and [Mog92] proposed a CMOS structure with local well contact for low temperature operation. However, we studied the method of minimizing energy-per-cycle by applying FBB asymmetrically to obtain the optimal trade-off between energy efficiency and performance advantages.

Figure 4.1(a) illustrates the pMOS in n-well and common p-substrate of nMOS devices which are connected to \( V_{dd} \) and \( \text{gnd} \) terminals respectively, to keep the parasitic diodes in reverse-bias mode. Figure 4.1(b & c) represents the symmetric configuration with fixed optimal and swapped FBB which is applied to both the pull-up and pull-down logic network respectively. In case of asymmetric configuration, the shortened well contact is connected either to the substrate voltage \( (V_{AFB} = 0V) \) called \textit{n-well forward bias} and the supply voltage \( (V_{AFB} = V_{dd}) \) called \textit{p-well forward bias} respectively as shown in Fig. 4.2(b & c). Thus, the applied bias voltage lowers the threshold voltage of either pull-up or pull-down logic network and it correspondingly increases the low-to-high or high-to-low transition rate respectively.
Figure 4.2: (a) Asymmetric configuration with single bias voltage (b) and (c) n-well and p-well forward bias with well-contacts connected to $V_{ss}$ and $V_{dd}$ rails respectively.

At the same time, the respective high-threshold voltage of pull-down and pull-up logic network maintains the static leakage current under control. This shows that applying FBB asymmetrically to the static CMOS logic improves performance by making one of the transition faster than the other with lower static leakage current in the ultra-low voltage regime, when noise margin and transition symmetry are of less significance.

To further extend our study, we simulated fan-out of 3 (FO3) inverter based 51-stage delay chain using NAND as control gate to enable oscillation as shown in Fig. (4.3). It consists of each equally sized inverters driving three other similar inverters and the delay chain oscillates by enabling the NAND gate as shown in Fig. (4.3). Since the bias voltage is same as that of the operating voltage, the fan-out of 3 inverter delay chain is simulated for wide range of supply voltages.

4.2.1 Power and Performance Comparison

Figure (4.4) illustrates the percentage of frequency increase with asymmetric configuration compared to conventional zero body bias scheme. The frequency improved by n-well forward bias ($V_{AFB} = 0V$) is lower than that of p-well forward bias ($V_{AFB} = V_{dd}$) due to lower driving capability of pull-up devices. It shows that in for-
4.2 Asymmetric Forward Body Bias - An overview

Figure 4.3: Fan-out (FO3) of 3 Inverter based 51-stage delay chain

Figure 4.4: The change in frequency of 51-stage delay chain with asymmetric bias voltage compared to conventional zero body bias scheme (all data points are normalized to conventional ZBB)
ward biasing only the pull-down network (p-well forward bias) is highly efficient in terms of performance advantages than the biasing only pull-up network (n-well forward bias). In addition, the maximum frequency increase is obtained at a particular bias voltage when the entire PUN and PDN is forward body biased symmetrically. For instance, in sub-threshold region ($V_{dd} = 0.2V$) and above, the maximum frequency is approached nearly at 0.1V and 0.2V respectively. It clearly suggests that, there exist an optimal bias voltage, at which the maximum performance is approached has chosen to be nearly half of the supply voltage ($V_{AFB} = V_{dd}/2$) and further increase will deteriorate its performance. Thus, the optimal bias voltage approaches maximum performance at nearly half of the supply voltage in the sub-threshold, near-threshold and above-threshold regime.

This result is in contradiction to our study with 13-stage adder based delay chain where the optimal point exists at nearly 0.1V asymmetric bias voltage. It shows that the bias voltage at which peak performance is reached depends on the logic depth and the circuit architecture. In case, the performance is of major concern, then one could introduce asymmetric FBB at the cost of power increase. However, from our study, the symmetric FBB would be beneficial to match the sub-threshold currents to attain optimal power and performance trade-offs.

Figure 4.5(a) shows the change in average active power with asymmetric configuration compared to conventional zero body bias scheme. It shows that the active power consumption increases with bias voltage and it reached to a maximum extent when it operates in the sub-threshold supply voltages ($V_{dd} = 0.2V$ & 0.3V). Although the p-well forward bias has larger performance improvement over its n-well counterpart, it has significant active power due to increased pull-down nMOS junction leakage current. The p-well forward bias has higher active power consumption compared to n-well forward bias in sub-threshold, near-threshold and above-threshold regions. It shows that there exist a clear trade-off between the active power and performance when the delay chain operates in n-well or p-well forward bias configuration. As discussed above, the optimal bias voltage is selected at nearly half of the supply voltage and further increase causes extra power without any performance advantages. Thus, the optimal bias voltage provides performance benefits with minimal active power increase compared to n-well and p-well forward bias configurations.
4.2 Asymmetric Forward Body Bias - An overview

Figure 4.5: The change in active and static power (average) of 51-stage delay chain with asymmetric bias voltage compared to conventional zero body bias scheme

The average static power of 51-stage delay chain is computed from the product of number of stages and average static power of FO3 inverter. Figure 4.5(b) shows that, increasing the bias voltage lowers the static power and it reaches lowest extent when it operates at supply voltage (p-well forward bias). Furthermore, the static power of n-well forward bias is higher than that of p-well forward bias, because forward biasing only the pull-up devices draws larger static current compared to the pull-down devices. In case of p-well forward bias, the zero body biased nature of pull-up logic resists the static current flow and hence it has lowered static power than its counterpart. Similar to our previous discussions, the optimal bias voltage is chosen at nearly half of the supply voltage \( V_{AFB} = \frac{V_{dd}}{2} \) to obtain the good compromise between the performance benefits and minimal static power. Thus, it clearly shows that the applying FBB of \( \frac{V_{dd}}{2} \) symmetrically provides optimal trade-off between the power and performance at the expense of external circuit overhead to generate the bias voltage. On the other hand, the n-well and p-well forward bias with
4 Low-Voltage Fine-grained FBB

negligible layout overhead shows moderate performance improvement with minimal operating power and hence it is best suitable for energy efficient applications.

Table 4.1: Power and performance comparison of 2-input NAND and NOR with conventional Zero body bias, n-well Forward bias and p-well Forward bias respectively \((V_{dd} = 0.5V)\)

<table>
<thead>
<tr>
<th>Schemes</th>
<th>NAND</th>
<th>NOR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Delay (ns)</td>
<td>Active Power (nW)</td>
</tr>
<tr>
<td>Conv. ZBB</td>
<td>0.23</td>
<td>100.7</td>
</tr>
<tr>
<td>n-well Forward bias</td>
<td>0.21</td>
<td>104.9</td>
</tr>
<tr>
<td>p-well Forward bias</td>
<td><strong>0.19</strong></td>
<td>107.5</td>
</tr>
</tbody>
</table>

4.2.2 Performance and Energy-Per-Cycle Comparison

Figure (4.6) represents the active and static energy-per-cycle of FO3 inverter based 51-stage delay chain for n-well forward bias, p-well forward bias, the optimal bias voltage and conventional zero body bias schemes. It shows that the moderate performance improvement with lower active energy and higher static energy is obtained with n-well forward bias compared to all other schemes. On the other hand, p-well forward bias shows considerable performance improvement with higher active energy-per-cycle with relatively lower static energy compared to other schemes. As a result, the optimum between energy and performance occurs when it operates either in n-well forward bias or p-well forward bias configuration respectively. In case of the optimally chosen bias voltage \((V_{AFB} = V_{dd}/2)\), the active and static energy-per-cycle exists within the n-well and p-well forward bias boundaries with considerable performance. Thus, the optimally chosen bias voltage improves the energy efficiency by forward biasing the entire circuitry symmetrically at nearly half of the supply voltage. Furthermore, it suggests that conventional zero body bias is
4.2 Asymmetric Forward Body Bias - An overview

Figure 4.6: The change in active and static power (average) of 51-stage delay chain with asymmetric bias voltage compared to conventional zero body bias scheme

mostly likely suitable for lower energy-per-cycle at severe performance penalty. On the other hand, the n-well and p-well forward bias schemes can be placed selectively depending on the power and the performance constrained critical paths since both lies at the boundary regions.

Table (4.1) illustrates the simulation results of 2-input NAND and NOR gate using conventional ZBB, n-well forward bias and p-well forward bias respectively. The p-well forward bias minimizes the NAND delay and the respective energy-per-cycle by 17% and 9% respectively, on the other hand, the n-well forward bias minimizes the NOR delay and the corresponding energy-per-cycle by 24% and 14% respectively, when compared to the conventional scheme. This lower energy-per-cycle in the NAND(NOR) p-well(n-well) forward bias is mainly due to the larger reduction of delay obtained from serially stacked nMOS(pMOS) devices respectively. Thus, the delay reduction due to forward biasing only the serially connected stacked devices is significant compared to forward biasing only the paralleled connected devices. From energy efficient view-point, the delay reduction should be large enough to overcome
the power increase in the forward bias configuration to obtain significant energy-per-cycle reduction. This shows that, the n-well and p-well forward bias could be used in the monotonic logic circuits to obtain highly energy efficient design.

Thus, the above discussion clearly suggests that the n-well and p-well forward bias configuration provides energy efficient design without much design complexity for ultra low voltage operation. Secondly, the optimum symmetrical \((V_{dd}/2)\) FBB improves the energy efficiency with minimal area overhead to generate forward bias voltage. Thus, this approach is not only suitable for the state-of-the-art static CMOS logic circuits, but it is highly applicable for precharge-evaluate logic circuits which is discussed in the next section.

### 4.3 Precharge-Evaluate Logic Circuits

Like dynamic circuits, the domino logic circuits has been a popular choice to obtain both low-power and high performance requirements. However, the continuous scaling of supply voltage limits its usage because of higher susceptible to circuit failure due to threshold voltage variation and noise injection. In-addition, it increases the subthreshold leakage current exponentially, which makes it even more susceptible to noise. It shows that domino logic does not scaled well like static circuits with the process technology and hence, it may not be a suitable for low-voltage low-power applications. To overcome the limitations of domino logic, a new noise-immune high performance logic style, called skewed logic[Sol00] or Monotonic Static(MS) CMOS logic[Tho99], has been proposed.

Skewed logic are fully complementary static circuits in which the pull-down network (PDN) and pull-up network (PUN) transistor widths are adjusted to make one of the transitions faster than the other (Figure (4.7)). Sizing the PDN and PUN to favor one of the transition faster is referred to as skewing. The same result is achieved with different supply voltages or transistors with different threshold voltages to speed up one of the transition. In this case, the transistor sizes need not be increased and hence the input capacitance and area are smaller at the expense of multiple power/ground lines or complex process technology. Skewed logic gates have performance comparable to that of dynamic circuits, whereas the noise toler-
4.3 Precharge-Evaluate Logic Circuits

Figure 4.7: Circuit topology and its connection (a) Clock-less Skewed NAND gate (Slow precharge) (b) Clock-less Skewed NOR gate (c) Clocked Skewed NAND gate (Fast precharge) [Skew Ratio = 3]

ance of skewed logic is better because it has no floating nodes. The floating node in a domino gate can be eliminated using a keeper device. However, the keeper cannot restore the correct state of the gate if a false transition occurs due to input glitch. Moreover, the skewed logic allows a trade-off between the gate delay and its noise margin. Because of higher noise tolerance, skewed logic is better than domino logic for high performance low voltage/low power applications. Like Domino gates, skewed gates should operate in two phases like precharge and evaluate modes in order to have monotonic transition at its gate output. During precharge, all the output nodes are precharged to the initial state $V_{dd}$ or $V_{ss}$ and the circuit performs the logic function during evaluation mode.

4.3.1 Topologies of Skewed Circuits

Skewed logic circuit falls in the category of precharge-evaluate logic families similar to dynamic logic circuits. Faster transitions are used for evaluation and slow transitions are used as precharge. The circuit topology of skewed logic gate is same as that of conventional static CMOS gate, except that the sizes of PUN and PDN devices are properly adjusted such that the gate favors a particular transition direction. Changing the ratio between PUN and PDN devices of the gate from its original value $W = \frac{W_p}{W_n}$ (Where $W_p$ and $W_n$ being the pMOS and nMOS transistor widths,
4 Low-Voltage Fine-grained FBB

Figure 4.8: Pulsed static CMOS - precharging the output node by input pattern without any clock scheme

respectively) to a new value \( W' = \frac{W_p}{W_n} \) to favor one transition direction is referred to as skewing and \( SR = \frac{W'}{W} \) called the skew ratio of the gate. Note that, the overall gate size and gate capacitance are invariant under skewing [Som99]. The circuit topology of skewed NAND and NOR gate without clock and skewed NAND with precharge device to make faster high-to-low transition and low-to-high transition are shown in the Fig. 4.7(a)(b) and (c), respectively. The logic gates with and without external precharge device are named clocked and clock-less skewed gates respectively. The skewed logic gates with faster low-to-high and high-to-low evaluation transitions are called high-skew and low-skew gates respectively.

4.3.2 Skew-Circuit Operation

Skewed logic circuits can achieve performance comparable to that of domino logic circuits [Kra82] when it operates in precharge-evaluate mode. A skewed gate is precharged to the logic value that allows only fast transitions during evaluation phase. For faster evaluation transition, the low-skew gates must be followed by high-skew gates, and vice versa. Precharging can be accomplished either by clocked skewed logic gates which precharge just like domino logic gates or by the propagation of precharged logic values through the logic chain originating from a clocked gate. For example, Figure (4.8) shows pulsed static CMOS logic in which the output node is precharged by the input pattern without any clock scheme to make faster evaluation transition. On the other hand, it is important that the precharging of
skewed logic gates between two clocked gates does not exceed the precharge phase of the clock period. [Tho99] considered clocking of all skewed gates, whereas [Sir01] and [Sol00] explored selectively clocking the skewed logic gates are shown in Fig. (4.9). Alternatively, precharging through propagation allows a smaller number of clocked gates. It leads to a lower clock load and hence potentially lower clock power consumption than domino circuits. Determining the selective clocking of skewed gates for low power operation is an integral step of the synthesis of skewed logic circuits.

4.3.3 Inverter Free Logic Synthesis

Conventional static CMOS logic is always synthesized using the flexibility of manipulating inverters in the logic network. The inverter-free constraint in static skewed logic design limits its flexibility, like domino logic gates. This constraint implies that all the logic inverters should be preformed at the clock phase boundaries, i.e., at the primary inputs or outputs where the inverters can be absorbed in registers. Thus the initial phase in skewed logic synthesis is to make the logic inverter-free. A straightforward approach is to convert the technology independent logic into AND, OR, and NOT gates only. Subsequently, the inverters can be propagated towards the inputs by applying simple De Morgan’s laws as shown in Fig. (4.10) stating at the primary
Figure 4.10: Propagating an inverter through logic gate (a) without fanout (b) with fanout

Figure 4.11: Forward propagation of intermediate inverters towards primary outputs
outputs. If an inverter is trapped at the fanout of a gate G, then gate G is duplicated for implementing both positive and negative phases and the inverter is pushed backward as shown in Fig. 4.10(b). Pushing the inverters back towards the primary inputs is guaranteed to restrict the increase in the size of the circuit to at most twice the original size and not increase the number of logic levels. This procedure transforms the given logic network into an inverter-free logic network with inverters at its primary inputs only.

In general, the logic duplication penalty for removing the trapped inverters all the way to the primary inputs can be quite heavy in terms of circuit area. [Ruc96] reports that an area overhead of 10% to 80% in microprocessor design partitions if all the trapped inverters are removed by pushing them to the primary inputs. Moreover, this area overhead also results in a substantial power dissipation penalty. One of the solutions is to propagate some of the inverters forward towards primary outputs thereby avoiding the logic duplication. Since some inverters can be propagated forward (Fig. 4.11) and others can be propagated backward (Fig. 4.12), the process may be very complex in a practical design and the choices grow exponentially. The process of propagating an inverter forward is equivalent to choosing the implementation phase (i.e., polarity) of the outputs so that the inverter is eliminated. This is illustrated in Figure (4.11), where $inv2$ is trapped at fanout net $N2$. Propagating $inv2$ back towards primary inputs will require duplication of fan-in logic cone of net $N2$. This duplication can be avoided by propagating $inv2$ forward towards primary outputs which is simply equivalent to implementing primary output $O2$ in negative phase. Since there are $2^n$ possible phase assignments for implementing $n$ primary outputs, it is non-trivial to find an optimal output phase assignment for minimum logic duplication. However, [Ruc96] addressed the problem of output phase assignment for minimum logic duplication in inverter-free dynamic logic synthesis by introducing an algorithm that significantly reduces the search space by pruning many fanout nets that must be considered for selecting an optimal output phase assignment. Thus, the above discussion shows that the skewed logic circuits are like dynamic circuits which is best suited for energy efficient design due to higher noise immunity and static nature which is highly desirable for ultra-low voltage applications. However, the clock power overhead could be reduced to certain extent
Figure 4.12: Backward propagation of inverters to obtain inverter-free logic
(a) logic with NAND/NOR/AND/OR/NOT gates (b) logic with AND/OR/NOT gates (c) propagating inverters back towards primary inputs without any logic duplication (d) combining the inverters trapped at intermediate fanouts (e) inverter-free logic after propagating the inverters back towards primary inputs with logic duplication.
by precharging the output node by input patterns or by using selective clocking scheme. While the area overhead to make inverter free logic is inevitable.

4.4 Precharge-Node based Forward Body Bias

In conventional FBB, applying constant body-to-source voltage to the entire static CMOS circuits increases the active leakage power from non-evaluation logic devices. For example, in simple 2-input NAND gate of input vector "11" or "00", only half of the logic devices are turned 'ON' for evaluation and the remaining devices reside in the steady state (turned 'OFF'). If the $V_{th}$ of logic devices is lowered by FBB, then the unavoidable active and static leakage power increase from non-evaluation devices. It shows that applying FBB to the entire circuit creates unwanted leakage power from the non-evaluation devices. Moreover, it is also impractical to place low-threshold and high-threshold transistors depending on the evaluation and non-evaluation transition. It shows the inadequacy of using constant body-to-source voltage in the conventional static CMOS circuits. To overcome these issues and to analyze the trade-off between energy-per-cycle and performance, we consider precharge-evaluate logic circuits. The detailed description of static skewed gates is given in the previous sections.

![Diagram of Precharge-Node based Forward Body Bias (PNFBB) technique](image)

Figure 4.13: Precharge-Node based Forward Body Bias (PNFBB) technique

To evaluate the method of selecting FBB for precharge-evaluate logic circuits, we proposed an self-adaptive fine-grained technique called *Precharge-Node based Forward Body Bias* (PNFBB), in which the forward body bias is applied to either pull-up or pull-down logic network based on the evaluation transition. (Figure 4.13).
The body terminal of the pull-down (pull-up) logic devices are connected to the output precharged initial state \( V_{dd} \) (GND). During the precharge phase, the output node \( V_{dd} \) (GND) biases the body of the pull-down (pull-up) devices and hence it reduces the threshold voltage to make faster high-to-low (low-to-high) transitions. During the evaluation phase, the body biased pull-down or pull-up logic network makes the faster evaluation transition or it remains in its present state depending on its input stimuli. However, body biasing the evaluation logic devices by the precharge node increases the sub-threshold current during precharge phase, but these increase is not significant compared to their performance advantages. This sub-threshold current starts decreasing, when the output node makes a high-to-low (low-to-high) evaluation transition. To summarize, there are two cases in which the sub-threshold current dominates: (1) when the output node remains in precharged state during evaluation phase and (2) when it is pre-charging or pre-discharging the output node to either \( V_{dd} \) or GND. The sub-threshold current arising from case (1) could be reduced provided when the output node makes high-to-low transition for each evaluation phase whereas the case (2) is inevitable. It shows that, the precharge node based intrinsic forward body bias technique increases the high-to-low or low-to-high evaluation transitions in ultra-low voltage region.

![Figure 4.14: High-speed static skewed inverter with PNFBB technique](image)

Figure 4.14: High-speed static skewed inverter with PNFBB technique

Figure (4.14) shows the high speed version of static skewed inverter where the body of both the pull-down network and evaluate devices are connected to output node to make faster high-to-low or low-to-high evaluation transition. Figure (4.15)
4.4 Precharge-Node based Forward Body Bias

Figure 4.15: **Low-leakage static skewed inverter with PNFBB technique**

shows the low-leakage version of static skewed inverter where only the PDN body is tied to the output node with high-threshold precharge and evaluate devices. In the latter case, the high-$V_{th}$ evaluate device suppresses the static leakage current compared to the high speed static skewed version. Since our major concern is on energy efficient design, we mainly focus on low-leakage static-skewed version.

4.4.1 Characteristics of PNFBB concept

(a) **Skew-ratio Vs PNFBB**

Static skewed circuits are sized to make either fast low-to-high or high-to-low evaluation transition, the efficiency of PUN and PDN forward body bias with transistor-sizing (skew ratio) is studied using fan-out of 4 inverter (FO4) delay. We simulated the FO4 inverter with equal sized buffers at the input and output driving stages to get the realistic signals using an industrial 130nm triple well process technology. The inverter under study is connected in PNFBB fashion and the transistor size is varied in steps of 0.5 skew ratio at 0.5 $V_{dd}$ to examine the evaluation transition rate.

Figure (4.16) shows the FO4 inverter delay with skew ratio, in which the solid lines represent forward body bias ($V_{bs} = V_{dd}$) and the dotted lines represent the conventional zero body bias respectively. The simulation result shows that PUN and PDN FBB reduces the evaluation delay by 23% and 12.4% respectively for minimum sized devices (skew ratio=1). As skew ratio increases, the percentage of evaluation delay
Figure 4.16: Fan-out of 4 inverter delay versus Skew ratio

decrease by FBB reduces from 23% to 12.5% for low-to-high and 12.4% to 2.6% for high-to-low transitions. Hence, it provides considerably lower input capacitance and diffusion area without much design complexity. This suggests that the forward body bias is highly efficient for smaller device compared to larger sized devices. Moreover, at iso-performance the PUN and PDN FBB provides 2X and 1X less skew ratio compared to conventional zero body bias.

(b) Output Load Capacitance Vs PNFBB

In this analysis, we assume a low skew 3-input NAND gate where PUN and PDN are sized such that the rise and fall delays are identical (skew ratio = 1). The output load capacitance of the gate can be divided into an intrinsic and extrinsic component, or

\[ C_L = C_{int} + C_{ext} \]

where \( C_{int} \) represents the intrinsic capacitance due to transistor sizing (diffusion and gate-drain overlap capacitance) and junction capacitance due to FBB, \( C_{ext} \) represents the extrinsic capacitance due to interconnect and fan-out loads. Figure (4.17) illustrates the low skew 3-input NAND gate with conventional zero body bias and the proposed method for difference output load capacitances and transistor sizing (skew ratio). As we know, increasing the transistor sizing ex-
Figure 4.17: Comparison of low skew 3-input NAND evaluation delay with conventional zero body bias and the proposed method for different output load capacitance and skew ratio.

ac erbates the diffusion capacitance and once it starts dominates the external load capacitance, there will be no speed improvement which is referred to as self-loading effect. The simulation results shows that, at symmetrical conditions of skew ratio = 1 the low skew PNFBB NAND3 gate has lower evaluation delay than the conventional ZBB for different output load capacitances. It means that the external output capacitance is high enough to hold the output node voltage to bias the pull-down network before the next transition arises. Hence, the larger evaluation delay reduction under symmetrical condition is highly suitable for driving larger fan-out gates. On the other hand, as skew ratio increased the efficiency of PNFBB is reduced due to increased intrinsic capacitance introduced by both FBB and diffusion capacitance by larger pull-down network. As we can see, the PNFBB low skew NAND3 gate shows better evaluation delay reduction as long as the intrinsic capacitance due to FBB and transistor sizing is lower than that of the extrinsic parasitic capacitance (i.e., $C_{int} < C_{ext}$). In contrast, there is no advantage in delay reduction when $C_{int} > C_{ext}$. Hence, it shows that the proposed method has lower delay reduction compared to that conventional scheme as long as the intrinsic capacitance is smaller than the extrinsic parasitic capacitance.
4.4.2 Simulation Results and Discussion

To validate the proposed method, we compared and analyzed the PNFBB simulation results with a dual-threshold voltage technique for slow and fast precharge static skewed and domino logic circuits [Jay05b]. In the dual-threshold voltage technique, the low and high threshold devices are placed either in the pull-up or pull-down logic paths respectively as shown in Fig. (4.18). In this, the low-threshold pull-up or pull-down devices indicated by grey shaded region make faster evaluation transition and the high-threshold pull-up or pull-down devices indicted by the unshaded region makes slower precharge transition. To make an initial study in terms of energy-per-cycle, evaluation delay and static power for PNFBB static skewed and domino circuits, the two cascaded two-input NAND gates forming an AND-OR structure is constructed as shown in Fig. (4.19.d). The design under test consists of equal sized input buffers for realistic input signals and it drives two equally sized NAND gate as a loading device. The basic circuit topology for the proposed technique and its connection to 2-input NAND gate with skew ratio of 3 for static skewed and domino gates are shown in Fig. 4.19.(a, b & c). In fast-precharge static skewed and domino logic gates, the output node is precharged by an additional precharge device for every clock transition. In contrary, the output node is precharged to its initial value by feeding a suitable input (for example, the input "00" precharges the
4.4 Precharge-Node based Forward Body Bias

Figure 4.19: (a) & (b) PNFBB 2-input skewed AND gate with static inverter (c) PNFBB 2-input Domino AND gate with static inverter (Skew Ratio = 3) and (d) Simulation setup with equal sized buffers at both sides of complex AOI logic.

NAND gate output to \( V_{dd} \) in slow precharge static skewed gate and it avoids the extra clock power required as discussed in the previous sections.

Figure (4.19) shows a PNFBB based slow, a fast precharge static skewed and a domino 2-input NAND gate, in which only high threshold logic devices are connected to the precharge node and the transistor width of the entire pull down network is varied to make a faster high-to-low evaluation transition. According to the non-inverting property, each low-skew gate should be followed by high-skew gate and vice versa. Hence, the faster low-to-high transition (high-skew) static inverter connected in PNFBB fashion are placed at the output of each low-skew NAND gate. The transistor width (skew ratio) of each gate is varied for different supply voltages to measure the evaluation delay and energy-per-cycle. The energy-per-cycle is measured by multiplying the precharge delay with active switching power and the static power is calculated for particular input vector (\( A=B=C=0 \)) during precharge phase.

Figure (4.20) illustrates that the proposed technique consumes lower energy-per-cycle compared to the dual-threshold voltage technique at relatively lower per-
4 Low-Voltage Fine-grained FBB

Figure 4.20: (a) Supply voltage versus Energy per cycle and (b) Supply voltage versus Evaluation Delay for PNFBB and Dual threshold voltage technique

formance in the near-threshold ($V_{dd} \approx V_{th}$) and above-threshold region ($V_{dd} > V_{th}$). Though the supply voltage is scaled down to reduce active switching power as shown in Fig. 4.21(a), the energy-per-cycle of the dual-threshold slow precharge skewed gate increases due to the static low threshold devices in the evaluation path. Furthermore, both dual-threshold domino and precharge skewed static gates consumes higher switching and static power compared to proposed technique (Figure 4.21(a)&(b)). The active switching and static power of PNFBB and the dual-threshold fast precharge skewed gates are relatively comparable. In PNFBB technique, the considerable sub-threshold leakage current reduction is achieved by allowing the logic devices to operate in high-threshold mode when it completes the evaluation and enters into next precharge transition. Thus, the PNFBB static skewed gate has relatively lower energy-per-cycle, moderate performance improvement and lower static power than the domino logic gates in near-threshold and above-threshold
4.4 Precharge-Node based Forward Body Bias

Figure 4.21: (a) Supply voltage versus active switching power and (b) Supply voltage versus static power for PNFBB and Dual threshold voltage technique

region. Furthermore, in both PNFBB and dual-threshold technique the functionality of slow and fast precharge skewed logic is preserved until 0.3V and 0.2V with full output voltage swing. In case of PNFBB and dual-threshold domino logic, the output voltage swing collapses below 0.3V and 0.4V respectively. Moreover, the static noise margin of dual threshold static skewed and domino logic gates degraded below 0.3V and 0.4V respectively and this might be due to the static low threshold evaluation devices. In our PNFBB technique, full output voltage swing up to 0.2V for both static skewed and domino gates are achieved and the circuit still operates at 0.1V with reduced output swing. Thus, the proposed approach is best suited for both static skewed and domino logic circuits in sub-threshold, near-threshold and above-threshold regions.

Static skewed AND-OR logic has higher noise tolerance than the domino logic for ultra-low voltage operation. To analyze this, the study of PNFBB fast precharge
static skewed AND-OR logic for various supply voltages and skew ratio is performed. The simulation results suggest that the energy-per-cycle increases as the supply voltage scale down due to larger precharge delay as shown in Figure (4.22). This result is in contradiction to our expectations that the scaling of supply voltage reduces the energy-per-cycle. This is not actually the case for precharge evaluate logic families, because increasing the skew ratio lowers the evaluation delay and correspondingly increases the precharge delay. This causes the energy-per-cycle to increase since it is a product of both active switching power and precharge delay. Though the switching power is decreased to reduce the energy-per-cycle, the significant increase in precharge delay substantially increases the energy-per-cycle at lower supply voltage. Hence, to optimally balance evaluation delay and energy-per-cycle, choosing an appropriate supply voltage is mandatory. In this case, operating at sub-threshold region consumes higher energy-per-cycle than the near-threshold region. Hence, operating at near-threshold (0.4V) with optimal skew ratio achieves best trade-off between energy-per-cycle and evaluation delay.

Table (4.2) illustrates the static skew fast precharge AOI gate for the proposed
Table 4.2: Comparison of fast precharge complex static skew AOI evaluation delay and energy-per-cycle with zero body bias scheme (skew ratio=3)

<table>
<thead>
<tr>
<th>Supply Voltage</th>
<th>Proposed technique</th>
<th></th>
<th></th>
<th>Dual-threshold voltage technique</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Evaluation Delay</td>
<td>Energy per cycle</td>
<td></td>
<td>Evaluation Delay</td>
<td>Energy per cycle</td>
</tr>
<tr>
<td>0.1V</td>
<td>6%</td>
<td>8%</td>
<td>37%</td>
<td>54%</td>
<td></td>
</tr>
<tr>
<td>0.2V</td>
<td>10%</td>
<td>11%</td>
<td>45%</td>
<td>44%</td>
<td></td>
</tr>
<tr>
<td>0.3V</td>
<td>10%</td>
<td>9%</td>
<td>42%</td>
<td>32%</td>
<td></td>
</tr>
<tr>
<td>0.4V</td>
<td>8%</td>
<td>8%</td>
<td>36%</td>
<td>21%</td>
<td></td>
</tr>
<tr>
<td>0.5V</td>
<td>5%</td>
<td>13%</td>
<td>28%</td>
<td>14%</td>
<td></td>
</tr>
<tr>
<td>0.6V</td>
<td>2%</td>
<td>13%</td>
<td>21%</td>
<td>14%</td>
<td></td>
</tr>
</tbody>
</table>

PNFBB and the dual-threshold technique with skew ratio=3. It shows that the proposed method improves the evaluation delay with minimal energy penalty whereas the dual-V<sub>th</sub> technique has larger evaluation delay reduction with significant energy increase. Thus, the PNFBB fast precharge static skewed logic is best suited for ultra-low voltage circuits than the slow precharge and domino logic gates. So far, in this section we discussed the precharge node based forward body bias scheme for ultra-low voltage precharge-evaluate logic circuits. The proposed technique is analyzed with dual-threshold voltage technique to study the energy-per-cycle and performance trade-offs. We found that the proposed method has lower energy-per-cycle, static power and higher noise immunity for both the static skew and domino AND-OR logic in lower supply voltage compared to the dual-threshold voltage technique. From our simulation results, we further conclude that a moderate performance improvement, lower energy-per-cycle is achieved with low-leakage PNFBB version when the static skewed and domino AND-OR logic operates in sub-threshold and near-threshold region.
4 Low-Voltage Fine-grained FBB
5 Simulation Results

In the recent years, the low-power requirements are getting momentum in battery driven VLSI systems. Since the power consumption of CMOS VLSI quadratically depends on the supply voltage, it causes the low-voltage circuits to be exploited. If VLSI circuits are operated in 0.5V to 0.8V supply voltage range for low power consumption, the threshold voltage of MOSFETs should be well below 0.5V to turn the MOSFETs. Threshold voltage between 0.1V and 0.2V causes 10nA-order subthreshold leakage current per logic gate in a standby mode, which leads to 10mA standby current for 1M-gate chips. This actually hinders the mobile applications powered by a small battery source. Hence, the precharge node based forward body bias technique has been presented in the previous chapter in terms of power and performance metric for $(V_{dd} <= 0.5V)$ low-voltage operation. In that, we showed the comparison between static and dynamic precharge-evaluate logic circuits with conventional zero body bias and the proposed FBB techniques. In the following section, we evaluate the static skewed 16-bit adder with modified carry look-ahead structure using conventional ZBB and PNFBB schemes. By nature, the carry-look ahead structure has intrinsic non-inverting behavior and hence its best suited for the mentioned precharge-evaluate logic circuits. Therefore, in this section we synthesize and study the 16-bit CLA adder using static precharge-evaluate logic circuits in terms of energy consumption and performance gain in the ultra-low voltage region.

5.1 Carry-Lookahead adder

When designing high performance adders, it is important to overcome the carry rippling effect that is still present in one form or another in both the carry-bypass and carry-select adders. The linear growth of adder carry-delay with the size of the input word for an n-bit adder may be improved by calculating the carries to each
5 Simulation Results

stage in parallel. The carry of the \( i \)th stage, \( C_i \), may be expressed as,

\[
C_i = G_i + P_i C_{i-1}
\]  

(5.1)

where \( G_i = A_i B_i \) is the Generate signal
and \( P_i = A_i + B_i \) is the Propagate signal
Expanding this yields

\[
C_i = G_i + P_i G_{i-1} + P_i P_{i-1} G_{i-2} + \ldots + P_i \ldots P_1 C_0
\]  

(5.2)

The sum \( S_i \) is generated by

\[
S_i = C_{i-1} \oplus A_i \oplus B_i
\]  

(5.3)

or

\[
S_i = C_{i-1} \oplus P_i
\]  

(5.4)

The size and fan-in of the gates needed to implement this carry-lookahead scheme can clearly get out of hand. As a result, the number of stages of look-ahead is usually limited to about four. For four stage of look-ahead, the appropriate terms are

\[
C_0 = G_0 + P_0.CI
\]

\[
C_1 = G_1 + P_1.G_0 + P_1.P_0.CI
\]

\[
C_2 = G_2 + P_2.G_1 + P_2.P_1.G_0 + P_2.P_1.P_0.CI
\]

\[
\]

The above expression represents the generic carry-lookahead adder [Wes93]. In general, fan-in limits carry look-ahead to groups of four bits. Because of this, multi-level look-ahead structures are used for longer word size. If we consider conventional CLA, each carry block generates three local carry-out terms from the \( P,G \) and carry-in terms. In addition, the \( PG \)-block generates the 4-bit group \( P,G \) terms and it has three levels of delay to the most significant carry-out terms. Thus, the critical path for the 64-bit conventional adder can be written as: \( A,B \Rightarrow G_0 \Rightarrow G_{3:0} \Rightarrow G_{15:0} \Rightarrow C_{16} \Rightarrow C_{32} \Rightarrow C_{48} \Rightarrow C_{52} \Rightarrow C_{56} \Rightarrow C_{60} \Rightarrow C_{61} \Rightarrow C_{62} \Rightarrow S_{63} \)
Figure 5.1: Generic block diagram of Carry-Look Ahead structure (a) Conventional CLA structure (b) Modified CLA structure

Figure 5.1(a) shows how the conventional longest critical path and the serially stacked devices in the group and carry blocks reduces the performance of the adder. To overcome this, [Nai92] proposed 96-bit 4.5ns CMOS adder based on the modified CLA structure. In the modified structure, the PG-block generates intermediate group 'P' terms ($P_{2:0}, P_{1:0}$) and group 'G' terms ($G_{2:0}, G_{1:0}$), in addition to the 4-bit group P,G terms. A Manchester carry chain structure is used in the generation of the group 'G' terms to reduce the transistor count. The carry-block generates three local carry-out terms ($C_1, C_2, C_3$). The generation of the intermediate group P,G terms means the delay to any carry-out term is only one level. Figure 5.1(b) shows the critical path of the modified CLA structure and this can be expressed as follows:

$$A, B \Rightarrow G_0 \Rightarrow G_{3:0} \Rightarrow G_{15:0} \Rightarrow G_{47:0} \Rightarrow C_{48} \Rightarrow C_{60} \Rightarrow C_{63} \Rightarrow S_{63}$$

The above expression shows the reduced critical path of the adder using modified CLA structure and hence it has higher performance than the conventional imple-
Figure 5.2: Gate-level implementation of modified static skewed 4-bit group PG and carry blocks using dynamic Manchester carry chain implementation. These high-speed is due to the availability of the intermediate group P,G terms at every PG-block, which enabled the most significant carry-out term delay of every carry-block to be a one-level delay. The PG-block of the modified CLA design has 41 transistors as opposed to 19 transistors for the conventional CLA design. However, this will minimally impact area because these transistors arise in the intermediate bits, which will not be a limiting factor in determining the bit cell size, which is discussed in detail [Nai92].

5.2 Design and implementation

5.2.1 PG and Carry Blocks

As discussed in the previous sections, the implicit non-inverting behavior of CLA structure is chosen to realize precharge-evaluate logic circuits without much de-
Figure 5.3: Transistor level circuit implementation of modified static skew 4bit group PG and carry blocks using dynamic Manchester carry chain (optional clock shown by dotted lines). Note that low-skew NAND gates and high-skew inverters are connected in PNFBB approach.
sign complexity. The gate level implementation of PG-block used in the modified CLA structure is realized using precharge-evaluate low-skew and high-skew gates are shown in Fig. (5.2). The PG-block generates intermediate group 'P' terms ($P_2:0, P_1:0$) and group G terms ($G_2:0, G_1:0$) in addition to the 4-bit group 'P' and 'G' terms. The circuit level representation of dynamic and static Manchester carry chain structure in the generation of group 'G' terms to minimize the transistor count are shown in Fig. (5.3) and (5.4). The PG block consists of 3 low-skew NAND gates with fan-ins 2, 3, and 4 followed by high-skew inverters. In addition, it consists of 3 dynamic or static Manchester carry logic gates. If carry generate (A.B) is true, then the output node discharges. If carry propagate (A+B) is true, then a previous carry may be coupled to the output node, conditionally discharging it. On the other hand, the static carry chain requires 'P' to be generated as $A \oplus B$. These NAO gates are followed by high-skew inverters. The carry block consists of three NAO gates followed by high-skew inverters, similar to PG block. In general, the clock network is used in the design of precharge-evaluate PG and carry blocks to increase the precharge transition delay. In fact, this extra precharge device (reset or clock) are optional, as discussed in the previous section. This extra precharge device and the clock network increases the area and power overhead, which is not desirable for low power operation. However, this will be a best choice for high speed operation. Since our focus is on energy efficient design, we evaluate both clocked and clockless precharge-evaluate PG and carry blocks which are shown in the Fig. (5.3) and (5.4). Further, we compare the results of precharge-evaluate static CMOS realization with static and dynamic Manchester carry chain at different skew ratios. The basic static skew CLA version with skew ratio=1 and the conventional zero body bias are compared with different schemes.

5.2.2 Simulation setup

The industrial 130nm high speed device characteristics are listed in the Table (5.1) and the spectre circuit simulator is extensively used throughout our analysis. Each static skew adder is experimented with 130nm triple well process technology using cadence virtuoso schematic editor and analog artist environment. The design under test is placed between the input buffers and the output load stages to set up a
5.2 Design and implementation

Figure 5.4: Modified circuit implementation of static skew 4bit group PG and carry blocks using static manchester carry chain. The low-skew and high skew gates are connected in PNFBB technique.

Table 5.1: UMC 130nm technology high speed device parameters (W/L=10/0.12)

<table>
<thead>
<tr>
<th>Device Parameters</th>
<th>HS nMOS</th>
<th>HS pMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{TON}$ (V)</td>
<td>0.38V</td>
<td>-0.33V</td>
</tr>
<tr>
<td>$I_D$ (@$V_{DS}=V_{GS}=1.2V$)</td>
<td>630nA/um</td>
<td>-270nA/um</td>
</tr>
<tr>
<td>$I_{OFF}$ (@$V_{DS}=1.2V, V_{GS}=0V$)</td>
<td>4nA/um</td>
<td>-4 nA/um</td>
</tr>
<tr>
<td>$V_t$ Body effect</td>
<td>0.094V</td>
<td>0.157V</td>
</tr>
</tbody>
</table>
Figure 5.5: Experimental setup of 16-bit static skew adder with modified CLA with input buffers and output load stages

realistic simulation environment, which is shown in Fig. (5.5). Each 16-bit adder consists of 4 slices of 4-bit adder and a carry look ahead logic to generate SUM and intermediate CARRY terms respectively. Each 4-bit adder generates the SUM bits and the intermediate PG terms for CARRY generation which is shown in Fig. (5.6). The optional clock runs both into the 4-bit adder slices and the CLA stages. Since there is no extra circuitry required to make CLA monotonic, this option shows the adder with negligible area and power overhead which will be proved later with simulation results. Each adder is configured to PNFBB and conventional zero body bias schemes for different skew ratios.

Our experiment is to determine the energy efficient adder by estimating the power and performance with the critical CLA block implemented with PNFBB and conventional zero body bias Manchester carry chain respectively. The critical paths are implemented with static skew dynamic and static Manchester carry chain, low-skew NAND gates and high-skew inverters with different skew ratios. We measured the precharge and evaluation propagation delay between the SUM output and CIN input when the input vector $A=FFFF$, $B=0000$ and the carry input makes low to high transition for all the adder considered. This pattern propagates the input
5.2 Design and implementation

Figure 5.6: Schematic cross-section of 16-bit static skew adder with modified CLA structure. The optional clock to precharge the output node is shown in solid dotted lines.

carry through all the CLA stages and finally makes the low-to-high carry output transitions. For example, the simulated waveform of 4-bit CLA with the proposed technique is shown in Fig. (5.7) with input patterns as explained above. Similarly, the active and static power is measured from the wide range of random input patterns and the respective steady state values.

5.2.3 Simulation Results and Discussion

As we discussed in the previous chapter, the domino logic circuits are not a suitable choice for low-voltage applications because of numerous difficulties like severe noise immunity and charge leakage. Static skew logic circuit nearly overcomes these bottlenecks and provide substantial advantages at the expense of logic duplication like domino circuits. Thus, static skew circuits can potentially provide a reasonable alternative for low voltage operation by replacing the domino logic circuits. As we know, reducing the evaluation transition delay in one direction increases the precharge delay in other direction. In such cases, the clocking precharge device might be needed to alleviate this adverse effect with larger design complexity to make sure
Figure 5.7: Simulated waveform of sub-block 4-b CLA of 16-b skew CLA adder using proposed PNFBB technique.
circuit performs correct operation. Adding precharge device to lower the precharge delay without evaluate device might cause incorrect circuit operation. Thus, in our case we considered both the precharge and evaluate device to improve the total delay to avoid incorrect circuit operation. Moreover, for heavily skew device clocking the precharge device improves the precharge delay significantly which reduces the overall cycle time with larger power budget.

Figure 5.8: Static skew 16-b CLA adder using dynamic manchester carry chain using clocking and non-clocking scheme for conventional zero body bias and the proposed technique ($V_{dd} = 0.5V$)

Figure (5.8) shows the precharge delay, evaluate delay, active and static power of 16-bit static skew CLA with critical paths implemented with PNFBB and conventional zero body bias schemes. As we can see, the evaluation delay decreases and the respective precharge delay increases with skew ratio. In the conventional ZBB
scheme, the evaluation delay reduction with skew ratio diminishes beyond SR=3 due to self-loading effect. Moreover, larger skew ratio of either pull-up or pull-down network exacerbates the precharge delay dramatically because of increased diffusion capacitance with minimal impact on evaluation delay reduction. From the power perspective, the highly stacked evaluate device in the CLA critical path lowers both the active and static power compared to the adder with non-clocking scheme. Apart from the clock power, the logic power consumed by the clocking static skew adder is lower than that of the non-clocking scheme. The significance of clock overhead depends on the size of the precharge and evaluate device and the number of clocked devices. By inserting minimum sized precharge and evaluate devices, one can attain lower evaluation delay at reasonable power increase. Thus, the clocking static skew CLA adder consumes lower logic power because of highly stacked devices present in the evaluation transition network compared to the non-clocking scheme.

As we see in Fig.(5.9), the total delay of static skew 16-b CLA adder can be
improved by 15% and 14% at skew ratio=1 in the non-clocking and clocking PNFBB scheme respectively. In contrast, the PNFBB technique increases the active power minimally by 3% and 4% in the non-clocking and clocking scheme respectively. Further, it depicts that the effectiveness of PNFBB technique is minimized beyond a skew ratio=3 for an static skew 16-b CLA adder with an insertion of precharge and evaluate devices. The highly stacked gates and the corresponding larger diffusion capacitance lowers the efficiency of evaluation delay reduction in the clocking static skew 16-b CLA adder. On the other hand, this highly stacked clocking scheme consumes less active and static power compared to non-clocking schemes (excluding clock power). Thus, the overall power increase is minimal in PNFBB technique because the PDN or PUN switched back to zero bias mode once the evaluation transition is finished. This selective forward biasing allows lower power consumption compared to conventional fixed forward body bias approach. This corresponds to lower active energy-per-cycle with significant evaluation delay reduction as shown in Fig. (5.9). At skew ratio=1, 19% and 16% evaluation delay reduction causes the 15% and 10% active energy reduction for non-clocking and clocking schemes at 0.5V supply voltage. As we discussed earlier, larger diffusion capacitance due to higher skew ratio minimizes the advantage of forward body bias if we increase the skew ratio beyond 3. For iso-performance operation, one could get considerable energy reduction when the operating voltage is reduced for fixed performance target in ultra-low voltage regime.

Figure (5.10) represents the precharge and evaluate delay, active power of non-clocking version of static skew 16-b CLA adder with static Manchester carry chain using conventional zero body bias and the PNFBB schemes. The critical path is implemented with static Manchester carry chain without any precharge and evaluate devices which is shown in the Fig. (5.4). This adder overcomes both clock and extra area overhead but it requires careful timing closure to ensure proper circuit operation. All skew gates and the pull-down Manchester carry chain are connected in PNFBB configuration excluding transmission gates. As we discussed earlier, the maximum skew ratio is limited to 3 above which there is no advantage with PNFBB scheme. Furthermore, the evaluation delay reduction with PNFBB adder exacerbates by reducing the supply voltage. The evaluation delay reduction in static
skew 16-b CLA adder with static Manchester carry implementation is lower than that of adder with dynamic Manchester carry chain. This is due to the nMOS transmission gate in the static carry chain and it is not configured in PNFBB fashion, which actually limits the evaluation delay reduction. At skew ratio=1, the evaluation delay reduced about 4.5% and 3% at 0.5V and 0.6V supply voltage respectively. One can attain better delay reduction when all nMOS pull-down devices in the static carry chain are connected to the output precharge node (this is not the case here). However, the active power increases substantially with minimal advantage in the delay reduction using PNFBB technique.

Thus, in this section precharge-node based forward body bias technique has been studied using 16-bit adder using modified CLA structure for higher energy efficiency.
in ultra-low $V_{dd}$ regime. This proposed approach consumes lower energy-per-cycle than the conventional zero body bias due to the self-adaptive body bias approach which reduces the evaluation delay and the active power during evaluation transition and the precharge condition, respectively.
5 Simulation Results
6 Conclusion and Future Work

At first, it has been shown that application specific hardware for the low-level tasks and software for high-level tasks together with dynamic energy scaling reduce the power consumption at the architectural and circuit level respectively.

Next, we have addressed a co-design approach in which the leakage reduction techniques and state-of-the-art logic circuits are combined together to effectively minimize the leakage power during active and standby operation. Further, we analyzed the dual-threshold PMOS power gating with and without buffers in different static logic circuits. Our simulation results show that the source signal gating is not required for static CMOS logic gates irrespective of the gated or non-gated signals received from the previous blocks. In case of pass transistor logic gates, the insertion of source signal gating is mandatory to reduce the standby leakage power.

In addition, we showed that the dual-threshold voltage header with buffer scheme shows significant standby leakage power reduction at the expense of extra transistor count compared to dual-threshold voltage with and without header scheme.

To accomplish higher energy efficiency, the sub-threshold circuit design has emerged as promising approach for applications where performance is not a primary concern. However, the higher impact on process variations and the low-performance solution circumscribed the effectiveness of sub-threshold operation in the nanometer regime. To minimize this issues, we introduced the forward body bias in the sub-threshold region to achieve both energy efficiency and to mitigate the influence of process variation. First, we showed the optimum between energy-per-cycle and performance metric in the ultra-low voltage regime with the forward body bias technique. To analyze these metric further, we evaluated symmetric and asymmetric forward body bias configurations in the sub-threshold, near-threshold and above-threshold regimes. The minimum energy-per-cycle for symmetric and asymmetric
Conclusion and Future Work

Forward body bias schemes with novel 13-stage ripple adder based delay chain is experimented with commercial 130nm triple well process technology. We found that the minimum energy-per-cycle exists at sub-threshold region of 0.2V with substantial performance improvement, and this will vary depending on process fluctuations and the forward body bias configuration used. Thus, the sub-threshold circuit design provides ultra-low energy operation with significant performance penalty and large process variations in the sub-100nm regime. However, the introduction of forward body bias in the sub-threshold region minimizes the process fluctuations and improves the performance at the expense of minimal energy increase, which is highly energy efficient compared to conventional sub-threshold operation.

Finally, we experimented our novel inherent fine-grained self-adaptive forward body bias approach for the precharge-evaluate logic circuits. In this scheme, the forward body bias is applied to the high threshold voltage of either the pull-up or the pull-down logic network depending on the output evaluation transitions. This approach improves the energy efficiency with considerable performance gain at negligible area and power overhead in the ultra-low supply voltage regime. To validate the proposed method, the 16-bit look-ahead adder has been designed and experimented with static and dynamic Manchester carry chain structure. At skew ratio = 1, an delay improvement of 15% and 14% is obtained, at the cost of 3% and 4% active power consumption in the non-clocking and clocking 16-b static-skew CLA adder, respectively. The higher diffusion capacitance due to larger skew ratio minimizes the advantage of forward body bias if we increase the skew ratio beyond the optimal value (skew ratio = 3), and hence the minimum size device (skew ratio = 1) with the proposed PNFBB technique is highly beneficial in terms of both energy and performance metric.

Since the proposed method is perfectly matched for precharge-evaluate logic families, this would be highly desirable for memory designers to develop energy efficient row and column redundant circuits, which improves both yield and profitability. Further, this approach is not only suitable for state-of-the-art CMOS devices but it could be easily imported to advanced independent double-gate devices without much design complexity. Hence, it would be interesting to introduce this approach in asymmetric SRAM to improve the read cell stability without much area over-
head for ultra-low voltage operation. This custom design approach can be easily realized in the standard cell design methodology, once the self-adaptive technique is hard-wired in the standard cell gates.
6 Conclusion and Future Work
Bibliography


Bibliography


Bibliography


Bibliography


Bibliography


