School of Electrical, Electronic & Computer Engineering

Strained Silicon Technology for Low-Power High-Speed Circuit Applications

By

Hiran Ramakrishnan

Technical Report Series

NCL-EECE-MSD-TR-2008-140

July 2008

Contact: H.Ramakrishnan@ncl.ac.uk

NCL-EECE-MSD-TR-2008-140

School of Electrical, Electronic & Computer Engineering,
Merz Court,
Newcastle University,
Newcastle upon Tyne, NE1 7RU, UK.

http://async.org.uk/

ABSTRACT

One of the principal economic drivers for the semiconductor industry is high performance, low power applications for the portable electronics consumer market. Unfortunately, the power dissipation resulting from the use of conventional CMOS technology in this area is becoming a critical design issue. Supply voltage reduction has been the preferred technique for reducing power dissipation. However, the associated compromise comes in the form of a drastic increase in circuit delay. Silicon technology based contemporary circuit design research is struggling extremely hard to continue the 'scaling process' to satisfy the criteria of low static power dissipation, while maintaining high-speed operation to meet the demands of the market place demands. However, it is accepted that scaling cannot continue indefinitely and other approaches such as new materials and device structures need to be devised to circumvent the inevitable barriers to the reduction in device dimensions.

A potential candidate to meet the criteria of low static power dissipation while maintaining high-speed operation is strained silicon (strained-Si) due to its high current drive capability attained through band gap engineering. This thesis explores the applicability of this technology to low-power and high-speed digital and analogue designs by analyzing the power and performance characteristics of a range of circuits. Overall, it was found that strained-Si technology offered improved performance and power dissipation when compared with conventional bulk CMOS. For example, strained-Si devices could operate with supply voltages closer to the threshold voltage than bulk CMOS thus reducing the dynamic power dissipation in circuits. Furthermore, with reduced supply voltage the noise characteristics of the circuit were not significantly impaired. Also for the same power dissipation as a bulk CMOS circuit an improved performance level was achieved in strained-Si. These advantages can be realised without any major changes being required in the standard bulk CMOS process; this also includes the ability to vary the percentage strain in the channel thus permitting the device characteristics to be tailored to suit power/performance requirements for a given application which can be quite diverse.

As devices dimensions are scaled down, process variations can have a dramatic effect on the overall circuit performance. The impact of process variability on particular design parameters on several strained-Si circuits was investigated using two statistical methods, namely Design of Experiments (DOE) and Response Surface Modelling (RSM). These tools permitted the identification and modelling of those process parameters whose variation would have the greatest impact on circuit performance. In the case considered, strained-Si circuit was more robust to process variations than the standard Si implementation.

ACKNOWLEDGEMENT

There are many people I would like to thank for a huge variety of reasons.

Firstly, I would like to express my deep and sincere gratitude to my supervisor Dr. Gordon Russell. His wide knowledge and logical way of thinking have been of great value to me. Special thanks to Prof. Alex Yakovlev. Their understanding, encouragement and personal guidance have provided a good basis for the present thesis.

I would like to acknowledge the support I received from Dr. Sanatan Chattopadhyay from the Microelectronics Technology Group and Dr. Koushik Maharatna from Southampton University for the valuable guidance and suggestions.

I would like to thank Newcastle University for providing me with the scholarship to finish my thesis on time.

During this work, I have collaborated with many colleagues from Microelectronics Design and Technology Group for whom I have great regard, and I wish to extend my warmest thanks to all those who have helped me with my work.

I would like to thank all the members in the School of Electrical, Electronics and Computer Engineering of Newcastle University who helped me tackle different problems that I encountered during the period of my study.

Special thanks to my friends Dr. Raghav Menon and Sooraj Ranjit to name a few and all those Malayalee families who have given me a wonderful time at home without which my life in Newcastle would have been boring.

Finally, I express my special gratitude to my parents, brothers and their families, in-laws for their loving support. I owe my loving thanks to my wife Reena Nair. They have lost a lot due to my research abroad. Without their encouragement and understanding, it would have been impossible for me to finish this work. At last but the most important, I do not forget God Almighty, who has helped me throughout to the successful completion of this work.

List of Publications

1. **H. Ramakrishnan**, S. Shedabale, G. Russell, and A. Yakovlev, “Stacked Strained Silicon Transistors for Low-Power High-Performance Circuit Applications”, *IEEE ECTC08*, pp. 1793-1798, Florida, May. 2008.

2. **H. Ramakrishnan**, S. Shedabale, G. Russell, and A.Yakovlev, "Analysing the Effect of Process Variation to Reduce Parametric Yield Loss", *IEEE ICICDT 2008*, pp. 171-175, Grenoble, Jun. 2008.

3. **H. Ramakrishnan**, S. Shedabale, G. Russell and A. Yakovlev, “Statistical Variability Analysis of a Mutual Exclusion Element for Strained Silicon Processes”, *Workshop on Impact of Process Variability on Design and Test DATE’ 08*, Munich, Mar. 2008

4. **H. Ramakrishnan**, S. Shedabale, J. Zhou, G. Russell, and A. Yakovlev, "Variability analysis of a high performance strained silicon Jamb latch synchronizer", *19th UK Asynchronous Forum 2007*, London, Sept. 2007.

5. **H. Ramakrishnan**, K. Maharatna, S. Chattopadhyay, and A. Yakovlev, "Impact of strain on the design of low-power high-speed circuits", *IEEE ISCAS 2007*, pp. 1153-1156, New Orleans, May. 2007.

6. **H. Ramakrishnan**, S. Chattopadhyay, K. Maharatna, and A. Yakovlev, "Exploration of potential of strained silicon CMOS for low-power circuit design”, *Int. Conf. on Ultimate Integration on Silicon (ULIS)*, pp. 29-32, Mar. 2007.

7. **H. Ramakrishnan**, S. Chattopadhyay, A.Yakovlev, S. S. Dlay, and A.G. O’Neill, “Design of s-Silicon inverters for future VLSI applications”, *Int. Conf. on Materials and Advanced Technologies (ICMAT 2005)*, Jul. 2005.

8. S. Shedabale, **H. Ramakrishnan**, G. Russell, A. Yakovlev, and S. Chattopadhyay, "Statistical Modelling of the Variation in Advanced Process Technologies using a Multi-level Partitioned Response Surface Approach", *IET Circuits, Devices and Systems Journal*.

9. V. Haran, **H. Ramakrishnan**, and S. Chattopadyay, “Effect of Strained Silicon device parameters on universal gate performance”, *International Summer School on Physics of Functional Micro & Nanostructures*, Germany, 2006.

10. B. Halak, S. Shedabale, **H. Ramakrishnan**, A. Yakovlev and G. Russell, **“**The Impact of Variability on the Reliability of Long on-chip Interconnect in the Presence of Crosstalk”, *System Level Interconnect Prediction (SLIP 2008)*, Newcastle, pp. 65-72, Apr. 2008

11. Y. L. Tsang, S. Chattopadhyay, S. Uppal, E. Escobedo-Cousin, **H. Ramakrishnan,** S. H. Olsen, and A. G. O’Neill, **“**Modelling of the Threshold Voltage in Strained Si/ Si1−*x*Ge*x*/Si1−*y*Ge*y*(*x* ≥ *y*) CMOS Architectures”, *IEEE Trans. Electron Devices*, vol. 54, pp. 3040-3048, 2007.

CONTENTS

List of Figures

Chapter1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

Microelectronics can be considered as the field which has experienced the greatest advances of the last century [1]. The transistor fabrication process has evolved over 50 years into one of the most advanced and complicated manufacturing technologies. The progress in semiconductor manufacturing technology has been driven by market requirements for electronic systems with increased performance and functionality [2]. The mechanism with which the semiconductor industry addressed the continual challenges of market forces has been through the process of device scaling whereby transistor dimensions are reduced. The initial concept of Small Scale Integration (SSI) of chips with transistors numbering in tens started in early sixties; with an initial 10 micron gate length [4-7]. The minimum feature size gradually scaled to 0.15μm leading the industry to the concept of Very Large Scale Integration (VLSI) in the early eighties with the number of transistors in the range of hundreds of thousands. Moore's Law [8] describes an important trend in the circuit integration which states that that the number of transistors that can be placed on an integrated circuit doubles approximately every eighteen months.

Out of the two conventional transistor technologies, Metal Oxide Semiconductor (MOS) and bi-polar; MOS Field Effect Transistor (MOSFET) has overtaken the bi-

polar transistors as the basic building block of modern integrated circuits [9]. This is due to the advantages of MOS technology over bi-polar for example, cheaper manufacturing cost, fewer fabrication process steps, lesser chip area, more functionality per unit area, etc. Modern Integrated Circuits (IC) for example microprocessors, microcontrollers, DSP chips and chips for communications pack more than 1 billion transistors on a single IC. The end result of the dimensional scaling of transistors over the years helped the dramatic evolution of portable electronic applications in all areas. Figure 1.1 shows the increase in the density of transistors in Intel chips manufactured during 1970 to 2005 [10] and is well keeping with Moore's Law.

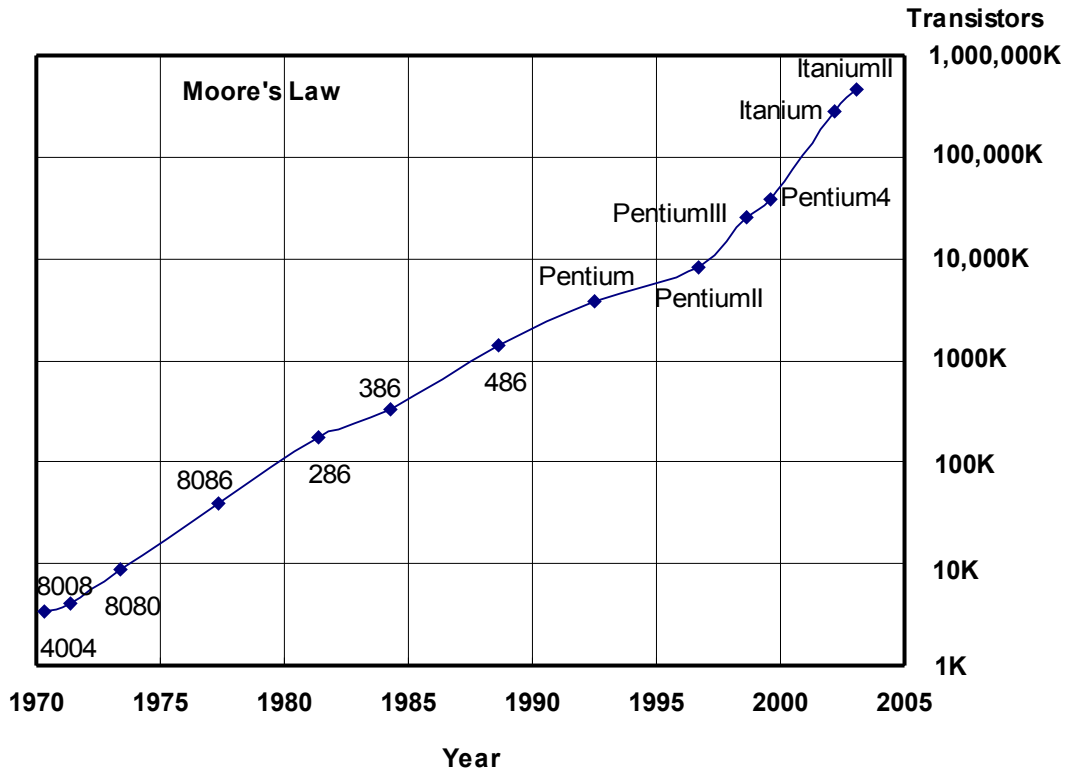


Fig. 1.1 The increase in the density of transistors in IC chips manufactured by Intel [10].

Figure 1.2 shows different stages of advancements achieved by microelectronic industry in terms of functionality per unit area. The first ever MOSFET transistor was designed by M. M. Atalla, D. Kahng, and E. Labate in late 1959 and marked as 1 in Figure 1.2 . Thereafter the main milestones in IC development are the first commercially available integrated circuit made by Fairchild Semiconductor in 1960, followed by the Intel 4004 microprocessor in 1971

which contains around 2400 transistors and then the Intel Pentium 4 released in late 90's containing tens of millions of transistors.

Although the reduction in transistor dimensions is the main objective when scaling devices, scaling also requires the reduction of other transistor parameters for example oxide thickness, junction depth, etc. so that the device delivers improved performance [9]. Device scaling can be realised in one of the two ways, either maintaining constant electric field within the device or constant supply voltage

[3-5, 7]. Constant field scaling yields the largest reduction in the power-delay product of a single transistor. However, it requires a reduction in the power supply voltage when the minimum feature size is decreased.

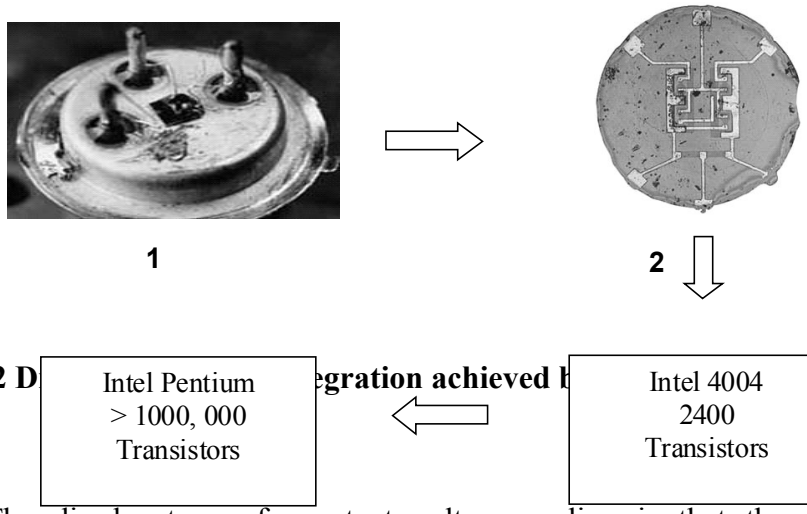


Fig. 1.2 D egration achieved b ndustry.

The disadvantage of constant voltage scaling is that the electric field increases as the minimum feature length is reduced. This leads to velocity saturation, mobility degradation, increased leakage currents and lower breakdown voltages. The scaling of MOSFET device parameters is illustrated by Table 1.1 where constant field, constant voltage and constant voltage scaling in the presence of velocity saturation are compared (α is the scaling factor). Another important aspect of transistor scaling is the scaling of parasitic resistances and capacitances. These

parasitic components do not necessarily scale with reduction of transistor dimensions [11].

Parameter	Symbol	Constant Field Scaling	Constant Voltage Scaling	Constant Voltage Scaling with Velocity Saturation
Gate length	L	1/α	1/α	1/α
Gate width	W	1/α	1/α	1/α
Field	ε	1	α	α
Oxide thickness	t_{ox}	1/α	1/α	1/α
Substrate doping	N_a	$α^2$	$α^2$	$α^2$
Gate capacitance	C_G	1/α	1/α	1/α
Oxide capacitance	C_{ox}	α	α	α
Transit time	t_r	$1/ α^2$	$1/ α^2$	1/α
Transit frequency	f_T	α	$α^2$	α
Voltage	V	1/α	1	1
Current	I	1/α	α	1
Power	P	$1/α^2$	α	1
Power-delay	P Δ t	$1/α^3$	1/α	1/α

Table. 1.1 Constant field scaling, constant voltage scaling and constant voltage scaling in the presence of velocity saturation is compared [7].

Scaling of transistor dimensions say, gate length and oxide thickness will improve the transconductance of the devices; this is an important metric which determines the speed of the devices [3-5, 10]. Furthermore, scaling reduces the area of the device which in turn increases the packing density resulting in an improved computational efficiency/unit area [2]. Scaled gate areas reduce the gate capacitance

which will contribute to faster switching speeds, lower dynamic energy consumption and higher throughput. In addition to these advantages for a given circuit, the increased packing density of transistors through scaling increases the number of die manufactured per slice hence increasing the production volumes.

Although the scaling process offered many advantages in the implementation of single chip digital systems regarding size, performance, cost and functionality, there is an adverse effect of the increase in on-chip static power dissipation as packing densities increased. However, with the introduction of CMOS technology in the year 1973, the problem of static power dissipation was largely removed as a logic gate realised with CMOS, ideally, only dissipates power when it is switching state compared to bipolar technology [3, 4, 12-14]. The comparison of physical characteristics of IBM systems based on bipolar and CMOS technologies are shown in Table 1.2 [14]. It can be seen from the table that the IBM G6 system, which is CMOS based offers more than double the performance of the fastest bipolar system (ES/9000). In addition it contains fewer components and has much lower space and power requirements [15].

IBM System	**ES/9000 9x2**	**S/390 G6**
Technology	**Bipolar**	**CMOS**
Total number of chips	5000	31
Power requirement (KVA)	153	5.5

Table. 1.2 Comparison of physical characteristics of bipolar and CMOS based IBM S/390 systems [14].

CMOS is also seen to a better alternative to bipolar transistors in terms of performance. As an example the performance trends of IBM S/390 systems based on CMOS and bipolar technologies are shown in Figure 1.3. The straight lines in Figure 1.3 represent the time-averaged exponential improvement in the performance of the technology [15]. The trend clearly justified the future and advantages of CMOS over bipolar technology [15].

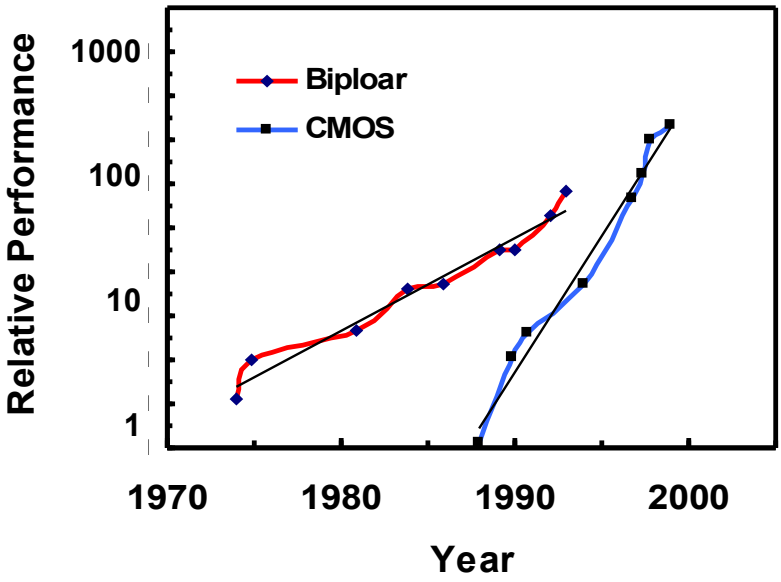


Fig. 1.3 Performance trends for IBM S/390 systems using bipolar and CMOS circuits [15].

Even though the introduction of CMOS technology to circuits reduces the static power dissipation by a considerable amount, the static power dissipation increases with scaling [10-13]. Figure 1.4 and Figure 1.5 show the chip power and frequency trends for Intel processors for various technology generations [10].

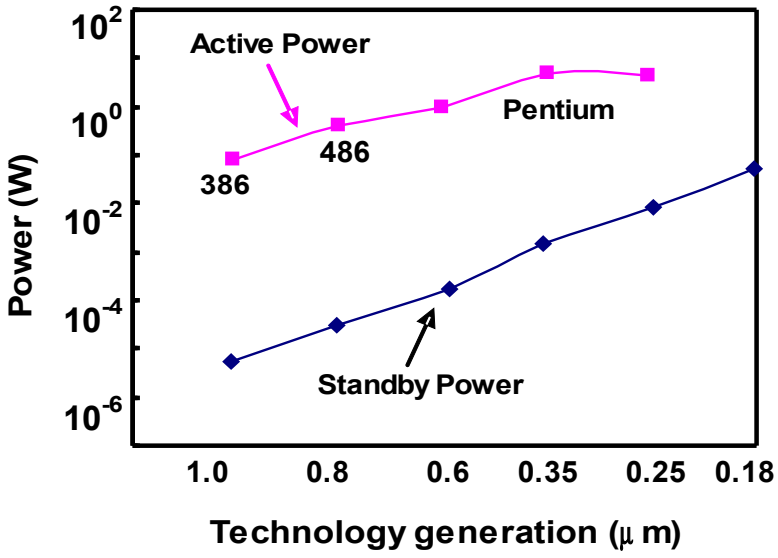


Fig. 1.4 Active and standby power trends for Intel process technologies [10].

It can be seen that the frequency trend is increasing. However, the standby and active power dissipation is also increasing linearly with technology scaling which is unacceptable. This is attributed to the fact that the scaling of oxide thickness, channel doping, channel length, gate length all contribute to higher leakage currents and subsequently to static power dissipation, for example, 100 million transistors each dissipating 0.01μA will lead to 1A of leakage current.

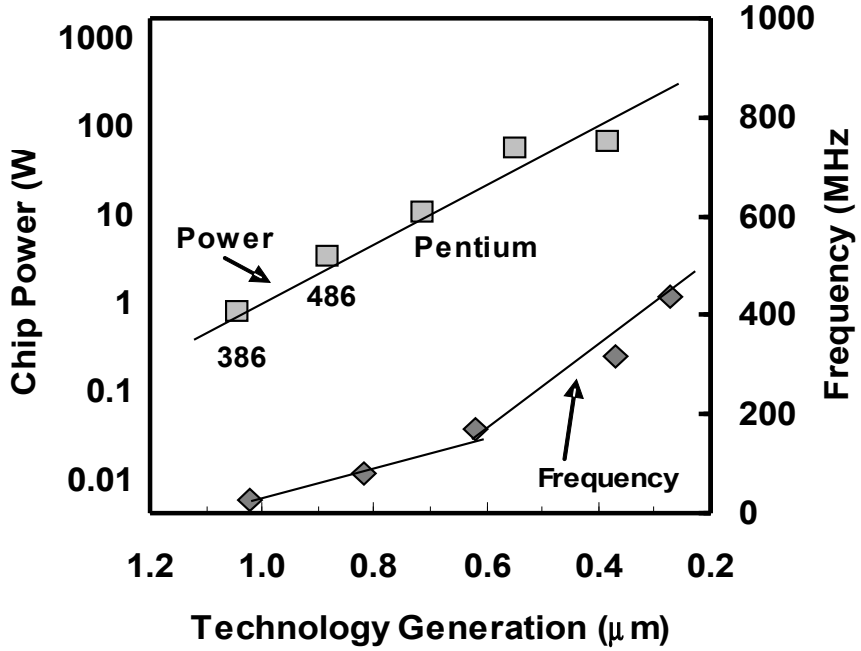


Fig. 1.5 Chip power and frequency trends for Intel process technologies [10].

The power dissipation issue of using conventional silicon based CMOS technology in digital circuit design is becoming a key design challenge to be addressed to meet the ever increasing demand for mobile computing with high data rate and the lack of significant advancement of battery technology [12]. Supply voltage scaling has been emerged as the preferable technique for power reduction owing to the quadratic relationship between supply voltage and active power dissipation [7]. However, the associated compromise comes in the form of drastic increase in circuit delay [7-13]. Although, for the applications such as wireless sensors circuit speed is typically overridden by requirement of energy efficiency, in

applications requiring higher performance voltage scaling poses a serious limitation in terms of speed [16]. The International Technology Roadmap for Semiconductors (ITRS) expects sub-1 V and 0.5 V nominal voltage supplies at 90 nm and 22 nm technology nodes, respectively, for low operating power [17]. Current circuit design research with Si-based CMOS is struggling extremely hard to satisfy the criteria of low static power dissipation while trying to maintain high-speed operation [17].

To date, silicon remains the material of choice for integrated circuit fabrication. Historically, this was mainly due to the ease with which the insulating silicon dioxide layer can be grown and its integrity maintained; silicon dioxide is an extremely good gate dielectric [3, 4]. Furthermore from an economic perspective due to the cost of replacing processing equipment and technologies, conventional device structures and materials will continue to be employed until it is no longer cost effective to do so. Enormous efforts are being made to extend the lifetime of current manufacturing technologies by continuing to adopt scaling techniques and developing new circuit design approaches such as, stacking of transistors, supply voltage scaling, asynchronous design methods, etc. to meet market place demands for even greater capabilities from electronic systems.

It is accepted that scaling cannot be continued indefinitely and other approaches need to be devised to circumvent the inevitable barriers to the reduction in device dimensions. Two approaches outlined below are currently being researched, namely the use of new materials and new device structures; although the microelectronics industry is reticent regarding their adoption, principally due to economic considerations (it costs over a billion US dollars for a fabrication facility to adopt to new manufacturing technologies) [18, 19].

1) New Materials:

The switching speed of devices can be enhanced by using materials which have higher carrier mobility than silicon. Subsequently a trade off between speed and power can be made by using supply voltage scaling. However, a radical departure from silicon is not likely to be accepted by the industrial community as the silicon technology offers low cost and high stability [3, 4, 10, 18, 19]. Recently, consideration has been given to evaluate the potential of new device materials for

example, strained silicon (strained-Si), Silicon Germanium on Insulator (SGOI) etc. The inherent superior transport property of strained-Si has already been exploited to fabricate high-speed low-threshold devices that can be used for the design and fabrication of high-speed circuits without any hidden cost. Accordingly, the use of strained-Si is now included in ITRS roadmap [17]. However, the device level performance improvement of strained-Si has not yet been exploited by the design community barring some experimental discrete RF circuits, which takes the advantage of high-speed operation of strained-Si transistors.

2) New Device Structures:

In recent years the potential of several new device structures have also been evaluated. Device structures for example, planar multigate, non-planar multigate (FINFET), tri-gate, Gate All Around Field Effect Transistors (GAAFET) have their own advantages and disadvantages when compared [20]. In a multigate device, the channel is surrounded by several gates on multiple surfaces, allowing more effective suppression of subthreshold leakage current. Multiple gates also allow enhanced drive current [18]. These advantages translate to lower static power consumption and enhanced device performance. Non-planar devices are also more compact than conventional planar transistors, enabling higher transistor packing density [18] to be achieved. The main challenges involved in the implementation of planar and non-planar device structures are with respect to lithography and patterning which in turn increases the cost of manufacturing.

1.2 Thesis Overview

The thesis explores the possibility of performance enhancement in terms of power and speed using strained-Si devices for digital circuit design. In this thesis, the general behaviour of digital CMOS circuits using strained-Si as the base material is studied. In the early chapters emphasis is given on noise performance analysis, speed and low power performance of circuits using strained-Si technology. In the analysis, experimental device data has been used to calibrate the simulator in the case of devices in the technology nodes ranging from 100nm to 300nm. For smaller devices ranging from 90nm down to 45nm ITRS [17] specifications provided by

MASTAR [21] for high performance, low power, and low standby power are taken into consideration.

Various design methodologies and modern IC process technologies to circumvent the scaling issues are studied. Silicon device characteristics are reviewed and compared and the device simulator MEDICI is calibrated with experimental data whenever and wherever possible. Material issues related to the strained-Si/relaxed SiGe, tensile and compressive strained device hetrostructures required for a CMOS technology are also discussed.

The work is performed by studying circuit characteristics of different basic circuit blocks subjected to different operating conditions and device dimensions. The work outlined in this thesis shows that even under the scaled supply voltage condition, strained-Si technology shows an improved performance compared to silicon technology. Furthermore, for the gate length considered (400nm to 45nm), strained-Si CMOS shows an improved performance over the conventional Si CMOS. Noise margin and metastability issues are crucial in the design space of any new technology to achieve desired functionality, and are found to be enhanced in the case of strained-Si technology compared to conventional Si technology [6, 24]. Furthermore the advantages associated with strained-Si technology are verified by introducing strained-Si technology to basic asynchronous circuit blocks for example Jamb Latch [25-26] and Mutual Exclusion Element (MUTEX) [27]. The choice of an asynchronous circuit block is obvious as the design technique is finding place in the modern day electronics as an alternative to conventional approaches due to its advantages which are discussed in detail in Chapter 2 [25-27].

Besides the solutions for performance enhancement and low leakage, variability is becoming problematic in the trend of scaling down to deep submicron (DSM) device dimensions. The transistor characteristics become more stochastic when the device dimensions shrink because the number of atoms which produce the transistor is reduced and hence variations in the process will be amplified. Besides the transistor characteristics, the effects of this statistical process variation on circuit properties become more observable. Furthermore as improved circuit performance is sought, this again proposes the requirement for the development of new technologies. To study the impact of variability in the manufacturing of transistors

the method of Design of Experiments (DOE) is introduced due to its computational efficiency over Monte Carlo Analysis which is a common method used in statistical analysis [28].

The entire work is executed by developing silicon and strained-Si technology library by simulation using TSUPREM4 (process simulator) [29], MEDICI (device simulator) [30], AURORA (parameter extraction program) [31], MINITAB (Design of experiments) [32] and PSPICE (Circuit Simulator) [33] subjected to different operating conditions, device dimensions and band gap engineered devices. The work outlined in this thesis shows that under scaled dimensions, different operating conditions, band-gap engineered devices and circuits based on strained-Si technology are beneficial for future circuit applications compared to the conventional silicon technology under the same conditions.

1.3 Thesis Organization

Having discussed the background and motivation for the work in this thesis the roadmap for the remainder of the thesis is outlined below.

Chapter 2 overviews characteristics of semiconductor materials and MOSFET devices which are used to build circuits. It gives an introduction to new semiconductor materials for example strained silicon (strained-Si) and Silicon On Insulator (SOI) and new structures for example FINFET and GAAFET. It also reviews different technological approaches adopted by the industry to improve the circuit performance. Conventional design methods to improve the power consumption criteria in circuits are outlined. Issues associated with scaling and variability on devices and their impact on high performance circuits are extensively reviewed.

Chapter 3 discusses the characteristics of strained silicon devices, comparing them to conventional silicon devices. An introduction to the TCAD tools used for simulation is given. A comparison of basic device characteristics (for example current/voltage characteristics, transconductance, short channel effects, threshold

voltages, etc) is given for both strained-Si and silicon devices. For the comparison of devices developed using both the technologies, similar processing and operating conditions are assumed. The parameters used for simulation are presented and discussed. The chapter also discusses the improved performance obtained for strained-Si devices, which can be used for building high performance circuits. Self-heating, is discussed here as a source variability.

Chapter 4 compares strained silicon CMOS and silicon CMOS inverter characteristics. Static and dynamic characteristics are given prime importance for the performance comparison. Ring oscillators made from strained silicon and silicon inverters are also measured for their frequency of operation through simulation using TCAD and experiments using semiconductor parameter analyzer. Calibration of the simulator with experimental device data is also discussed throughout the chapter when and where necessary. ITRS specifications provided by *MASTAR* are used for technologies ranging from 90nm down to 45nm due to the lack of experimental data available for these devices. This chapter also gives an insight into strained silicon based low power circuits and the advantage of using strained silicon with low power methodologies for Deep Sub Micron (DSM) circuits. The analysis of digital circuits using strained-Si technology in terms of power and speed are discussed. A deep insight into timing specifications of digital circuits and the advantages of using strained-Si is also presented.

Variability in IC manufacturing is presented in Chapter 5. A statistical method called Design of Experiments is used to analyze the process variability. Statistical screening method is performed to determine the most significant parameters. Basic asynchronous blocks for e.g., Jamb Latch, MUTEX, realized with strained-Si technology under different operating conditions are discussed. The impact of process and operational variability on the characteristics of these circuits are presented.

The conclusions and future work are presented in Chapter 6. The variability analysis performed here using the statistical technique DOE/RSM seems to be inefficient when large number of process parameters is considered due to the interaction of process parameters. New statistical technique need to be introduced to perform the variability analysis of deep sub-micron devices which are more prone to

process variability. The analysis needs to be extended to CMOS devices based on new materials and transistors based on new structures.

References

[1] C. A. Mead and L. Conway, *Introduction to VLSI systems*. Boston: Addison-Wesley, 1980.

[2] K. P. White, Jr., W. J. Trybula, and R. N. Athay, "Design for semiconductor manufacturing. Bibliography", *IEEE Transactions on Components, Packaging, and Manufacturing Technology, Part C,* vol. 20, pp. 73-86, 1997.

[3] S. M. Sze, *Physics of Semiconductor Devices*. New York: Wiley, 1981.

[4] M. Shur, *Physics of Semiconductor Devices*. New York: Prentice-Hall, 1990.

[5] R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*. New York: Wiley, 1977.

[6] J. F. Wakerly, *Digital design : Principles and Practices*. New Jersey: Prentice Hall, 1990.

[7] G. Baccarani, M. R. Wordeman, and R. H. Dennard, "Generalized scaling theory and its application to a ¼ micrometer MOSFET design", *IEEE Transactions on Electron Devices,* vol. 31, pp. 452-462, 1984.

[8] G. E. Moore, "Cramming More Components Onto Integrated Circuits", *Proceedings of the IEEE*, vol. 86, pp. 82-85, 1998.

[9] R. H. Dennard, F. H. Gaensslen, H.-N. Yu, V. L. A.-R. Rideout, V.L., et al., "Design Of Ion-implanted MOSFET's with Very Small Physical Dimensions", *Proceedings of the IEEE*, vol. 87, pp. 668-678, 1999.

[10]S. Thompson, P. Packan, M. Bohr "MOS Scaling: Transistor Challenges for the 21st Century," *Intel Tech. Journal Q3'98*. Portland Technology Development, Intel Corp.

[11]R. Gonzalez, B. M. Gordon, and M. A. Horowitz, "Supply and threshold voltage scaling for low power CMOS", *IEEE Journal of Solid-State Circuits,* vol. 32, pp. 1210-1216, 1997.

[12]G. Timp, A. Agarwal, F. H. Baumann, T. A.-B. Boone, T., et al., "Low leakage, ultra-thin gate oxides for extremely high performance sub-100 nm nMOSFETs", *International Electron Devices Meeting*, 1997

[13]A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-power CMOS digital design", *IEEE Journal of Solid-State Circuits,* vol. 27, pp. 473-484, 1992.

[14]S. G. Narendra and A. Chandrakasan, *Leakage in Nanometer CMOS Technologies (Series on Integrated Circuits and Systems)*. New York: , Springer-Verlag, 2005.

[15]R. D. Issac, "The Future of CMOS Technology", *IBM journal of Research and Development*, vol. 44, pp. 369-378, 2000.

[16]L. Nazhandali, B. Zhai, A. Olson, A. A.-R. Reeves, A., et al., "Energy optimization of subthreshold-voltage sensor network processors", *International Symposium on Computer Architecture*, 2005

[17]International Technology Roadmap for Semiconductors, "Process Integration, Devices and Structures", 2007 Ed.

[18]W. J. Dally and J. W. Poulton, *Digital Systems Engineering.* Cambridge: Cambridge University Press, 2000.

[19]K. R. Botkar, *Integrated Circuits.* Newdelhi: Khanna, 1991.

[20]H. S. Wong, "Beyond the conventional transistor", *IBM Journal of Research and Devlopment*, vol. 46, pp. 537-570, 2002.

[21]T. Skotnicki, C. Denat, P. Senn, G. Merckel and B. Hennion, "A New Analog/Digital Model for Sub-halfmicron MOSFETs", *International Electron Devices Meeting*, 1994.

[22]S. H. Olsen, K. S. K. Kwa, L. S. Driscoll, S. A.-C. Chattopadhyay, S., et al., "Design, fabrication and characterisation of strained Si/SiGe MOS transistors", *IEE Proceedings - Circuits, Devices and Systems,* vol. 151, pp. 431-437, 2004.

[23]Y. J. Yang, W. S. Ho, C. F. Huang, S. T. Chang, and C. W. Liu, "Electron mobility enhancement in strained-germanium n-channel metal-oxide-semiconductor field-effect transistors", *Applied Physics Letters*, vol. 91, 2007.

[24]L. Kleeman and A. Cantoni, "Metastable Behavior in Digital Systems", *IEEE Design & Test of Computers,* vol. 4, pp. 4-19, 1987.

[25]J. Sparsø and S. B. Furber, *Principles of asynchronous circuit design: a systems perspective*. Boston: Kluwer Academic Publishers, 2001.

[26]J. Zhou, D. Kinniment, G. Russell, and A. Yakovlev, "A robust synchronizer", *IEEE Computer Society Annual Symposium on Emerging VLSI Technologies and Architectures*, 2006.

[27]C. Dike and E. Burton, "Miller and noise effects in a synchronizing flip-flop", *IEEE Journal of Solid-State Circuits,* vol. 34, pp. 849-855, 1999.

[28]D. C. Montgomery, *Design and Analysis of Experiments*. New York: Wiley, 1991.

[29]TSUPREM4 User Manual: *Two-Dimensional Process Simulation Program*. Sunnyvale: Synopsys Inc, 2000.

[30]MEDICI User Manual: *Two-Dimensional Device Simulation Program*. Sunnyvale: Synopsys Inc, 2000.

[31]AURORA User Guide: *AURORA Simulation Program.* Sunnyvale: Synopsys Inc, 2000.

[32]B. F. Ryan, B. L. Joiner, T. A. Ryan, and T. A. Ryan, *Minitab Handbook*. Boston: Duxbury Press, 1985.

[33]P. W. Tuinenga, *SPICE: A Guide to Circuit Simulation and Analysis using PSpice*. New Jersey: Prentice Hall, 1988.

CHAPTER 2

DEVICES AND NEW MATERIALS FOR FUTURE CIRCUITS

2.1 Background

This chapter gives an introduction to semiconductor materials, devices and technology scaling. It also emphasizes the need for new technologies, structures and materials.

2.1.1 Insulators, Semiconductors and Conductors

A material may be categorized as an insulator, semiconductor or a conductor depending upon its conductivity which is a function of the number of carriers (number of electrons and/or holes) available to take part in conduction process [1, 2]. The band structure of the material permits the analysis of its conducting properties. A very poor conductor of electricity is called an insulator (approx 10^7 carriers/m^3), an excellent conductor is a metal (approx 10^{28} carriers/m^3); and a substance whose conductivity lies between an insulator and a conductor is a semiconductor [2].

2.1.2 Energy Bands

An energy band comprises large number of energy levels; the number and spacing of which depends on the atomic structure of the material and the "Pauli Exclusion" principle [3]. The energy bands of an insulator, semiconductor and conductor are shown schematically in the Figure 2.1. Insulators are characterized by having a large band gap. Since the electron is unlikely to acquire sufficient externally applied energy to carry it from the filled valence band into a vacant conduction band, conduction is negligible. Diamond is an example of an insulator [1, 2].

A substance for which the width of the forbidden energy (band gap) region is relatively small is called a semiconductor. The most common semiconductors are either germanium (Ge) or silicon (Si). These semiconductors have a band gap of 0.785eV and 1.21eV respectively. However, energies of this range cannot be acquired. Hence the conduction bands remain empty and valence band remains full at low temperature and hence semiconductors behave as insulators at low temperatures. When the temperature is increased the electrons in the valence band acquires thermal energy and hence will move to the conduction band. In the conduction band these electrons are free to move and hence will be influenced by a small amount of electric field and then this potential insulator starts conducting slightly. These are called intrinsic semiconductors [2, 3].

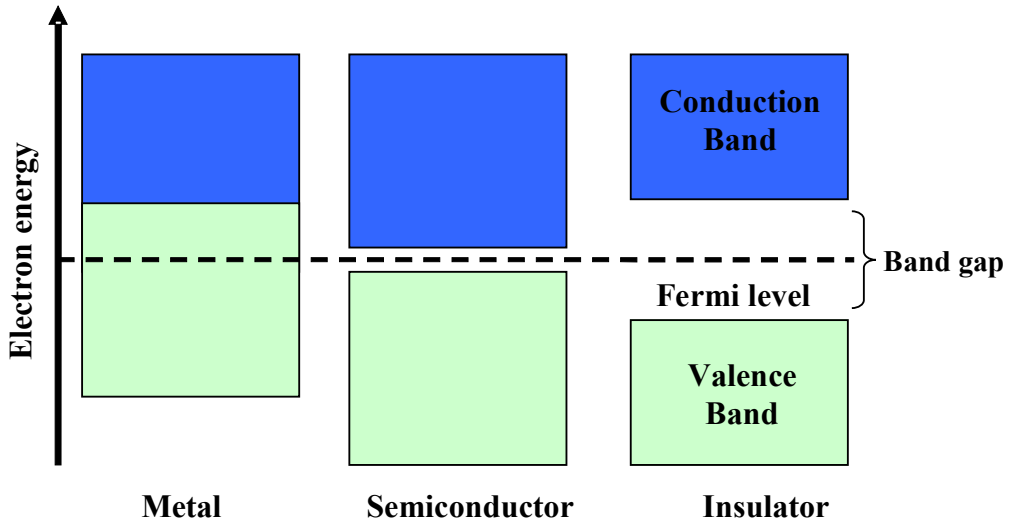


Fig. 2.1 Energy Bands.

The band structure of the material may have a merged valence and conduction band. In thermal equilibrium conditions electrons may acquire additional energy and move on to higher energy states. Since this movement constitutes a current this material is called a conductor [3].

In a Si lattice shown in Figure 2.2, all the atoms bond covalently to four neighbours, leaving few free electrons to conduct electric current. In this state Si is said to be intrinsic. The conductivity of an intrinsic semiconductor can be increased by introducing other materials into the structure, whereupon the semiconductor is said to be extrinsic. The materials which are added to the intrinsic semiconductor are called impurities and the mechanism of adding impurities to it is called doping [3]. Doping not only increases the conductivity of the semiconductor material, but also produces current carriers which are either predominantly holes (holes are actually the deficiency of electrons) or electrons. The extrinsic semiconductors are of two types: n-type and p-type. In an n-type extrinsic semiconductor, electrons are the majority carriers and in p-type, holes are the majority carriers [2].

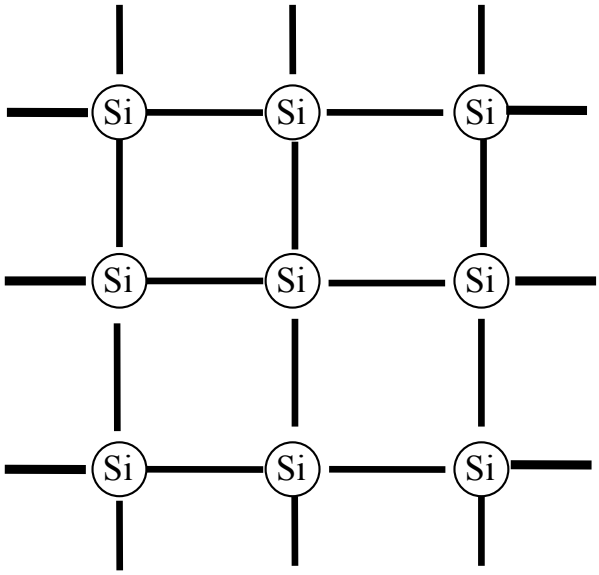


Fig. 2.2 Silicon Lattice.

In n-type doping, phosphorus (P) or arsenic (As) is added to the Si in small quantities. Phosphorus and arsenic each have five outer electrons; consequently four covalent bonds are made leaving the fifth electron free to move (donors) around in

the lattice. It takes only a very small quantity of the impurity to create enough free electrons to allow current to flow through the Si.

In p-type doping, boron (B) or gallium (Ga) are used as the dopant. Boron and gallium each have only three outer electrons. When introduced into the Si lattice, three covalent bonds are formed leaving a vacancy called 'hole' in the lattice where a Si electron has nothing to bond to. The absence of an electron creates the effect of a positive charge. A hole accepts an electron (acceptor) from a neighbour, moving the hole over a space [2].

2.1.3 Carrier Transport

Carriers within the semiconductor will move freely even without any externally applied force electric field because of the thermal energy associated with all the individual particles. At room temperature the thermal velocity of electrons in bulk semiconductors is about 10^7 cm/s [3]. In the absence of an applied electric field, the carrier motion is random and the carrier moves with high velocity through the semiconductor and frequently changes its direction. When an electric field is applied, the random motion still occurs however in addition there is an average net motion along the direction of the field. These effects are shown in Figure 2.3.

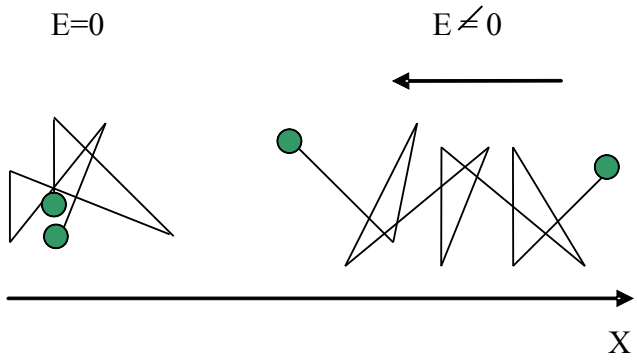


Fig. 2.3 Random motion of carriers in a semiconductor with and without an applied electric field.

When an electric field is applied to a semiconductor, due to the electrostatic force, the carrier first accelerates and then reaches a constant average velocity. This is due to the carriers being scattered as a result of the impurities present in the semiconductor and lattice vibrations [3]. The ratio of the velocity of carriers and the

applied electric field is called mobility [3]. Another very important scattering mechanism is the interaction of electron with a phonon (quantised mode of lattice vibrations) which contains the thermal energy of the crystal. When the density of phonons in the solid increases, the scattering time (time between collisions) decreases hence the mobility decreases. The density of phonons increases with increase in temperature and hence the mobility decreases with temperature. Impurities are efficient scattering centres especially when they have a net charge. Ionized donors and acceptors in a semiconductor are a common example of such impurities. The amount of scattering due to electrostatic forces between the carrier and the ionized impurity depends on the interaction time and the number of impurities. Larger impurity concentrations result in a lower mobility [3].

The velocity attained by a carrier by the application of an electric field is called the drift velocity and the current is called the drift current. There is another component of current called diffusion current. Diffusion is the process of particles distributing themselves from regions of high concentration to regions of low concentration [4]. If this process is left undisturbed, there will eventually be a uniform distribution of particles. Diffusion does not need external forces to act upon a group of particles. The difference between drift current and diffusion current is that drift current depends on the electric field applied; if there is no electric field, there is no drift current. Diffusion current occurs even though there is no applied electric field to the semiconductor [3] provided that a concentration gradient exists.

The electron and hole mobilities reduces when doping is increased [2]. Semiconductors with a low doping concentration have a constant mobility which is primarily limited by phonon scattering [2]. At higher doping concentrations the mobility decreases due to ionized impurity scattering with the ionized doping atoms. The actual mobility also depends on the type of dopant that is either n- or p-type. These effects are shown in Figure 2.4.

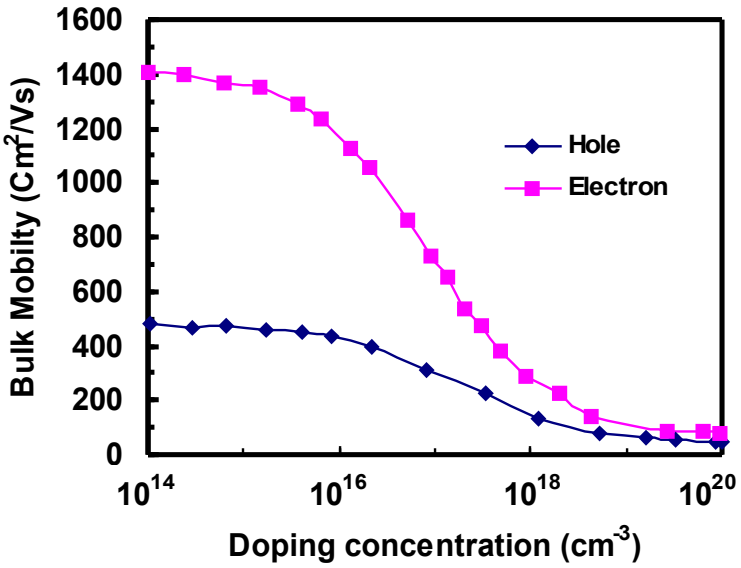


Fig. 2.4 Bulk electron and hole mobility versus doping density for Si [3].

Carrier mobility decreases with increase in temperature and its dependence is similar for both electrons and holes and is shown in Figure 2.5. The surface and interface mobility of carriers is affected by the nature of the adjacent layers or surface. This reduction of mobility finds its importance in modern hetro-structure semiconductor devices [2]. The presence of charged surface states further reduces the mobility in a similar way to ionized impurities [3].

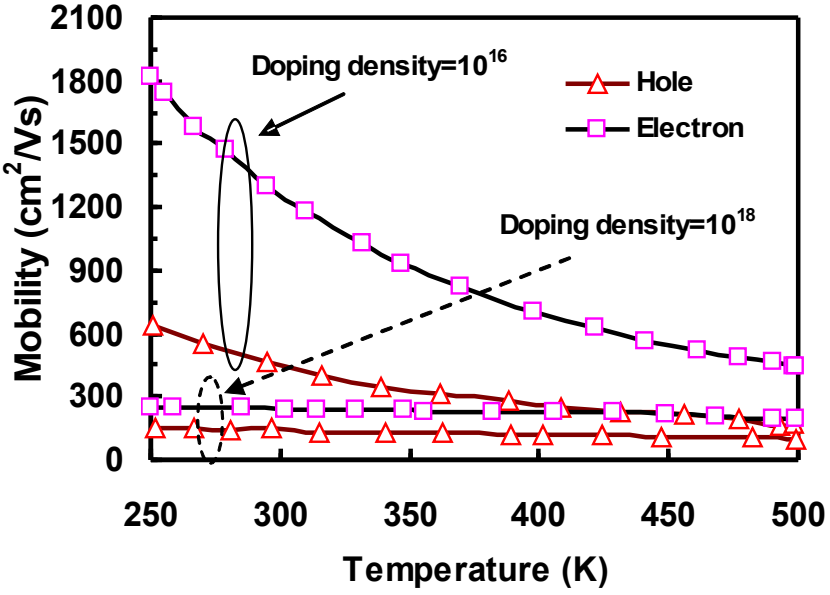


Fig. 2.5 Electron and hole mobility versus temperature [10].

2.1.4 Semiconductor Devices

By doping a semiconductor with p-type or n-type impurities its conductivity can be improved as discussed in the previous sections. A diode is the first and the simplest semiconductor device. It allows the current to flow only in one direction and hence it finds applications in for example, rectifiers. The Bipolar Junction Transistor (BJT) and Metal Oxide Semiconductor Field Effect Transistor (MOSFET) are the two different types of transistors [4]. The MOSFET has evolved as the transistor of choice for modern integrated circuits due to its advantages over BJTs for example, less area and low power consumption. The principle of operation of a MOSFET was first described by Lilienfield in his patent in 1926 [4, 5]. However, a successfully working MOSFET was not built until 1960 by Kahng and Atalla [2]. The MOS transistor has become the basic building block of most integrated circuits. Most modern complex ICs such as microprocessors, ICs used in graphics cards, and DSP chips, pack more than 100 million MOS transistors on a single chip. The microelectronics industry is striving towards the integration of billions of transistors onto a single chip. This ultra-high level of transistor integration is the result of transistor scaling over the past 40 years [2, 4]. Scaling of MOSFETS and its issues are discussed in detail in later chapters.

2.1.5 MOSFET and its Characteristics

A MOSFET is based on the modulation of charge concentration caused by a MOS capacitance. The structure of a MOSFET is shown in Figure 2.6. The MOSFET has two terminals, called source and drain, which are connected to highly doped regions which are separated by a region called the channel. These regions are either p- or n- type, but they must both be of the same type. The highly doped regions are denoted by a '+' following the type of doping as shown in the Figure 2.6 and are separated by a doped region of opposite type, known as the body or substrate. This region is not so highly doped. The third electrode in the MOSFET, called the gate, is located above the body and insulated from all of the other regions by an oxide (usually an oxide of Si) [2].

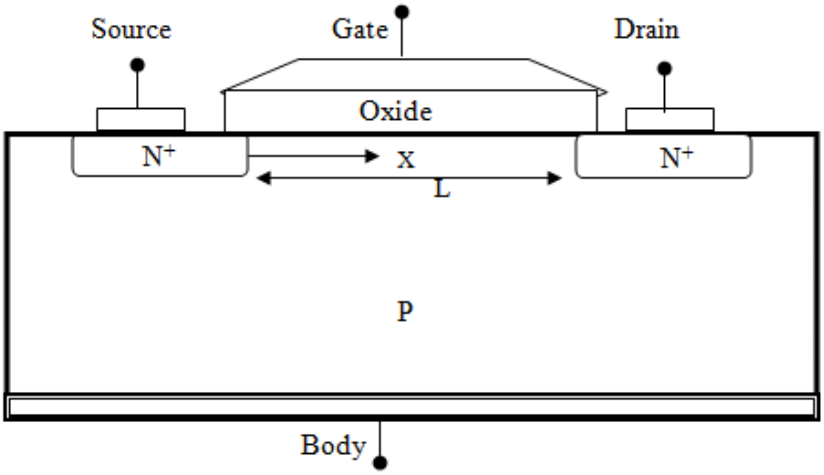


Fig. 2.6 Structure of a MOSFET.

The source is so named because it is the origin of the charge carriers (electrons for n-channel, holes for p-channel) that flow through the channel; similarly, the drain is where the charge carriers leave the channel. The MOSFET can be of n-channel or p-channel depending on the doping material in the source and drain of the MOSFET [4]. The MOSFET can be of two types; Depletion MOSFET and Enhancement MOSFET. In the case of depletion type of MOSFET the channel is lightly doped with the same material as that of source and drain to reduce the threshold voltage. The operation of both types of MOSFETs is similar and is used depending on the applications. Since the enhancement MOSFET is common, the operation discussed here are based on enhancement MOSFET only. The different modes of operation of an enhancement type MOSFET are discussed below [4].

Operating modes

In the case of n-channel MOSFET, when there is no voltage applied to the gate there is no channel formation between source and drain and hence there is no current flow between them. However, when a positive gate-source voltage is applied, it creates a channel at the surface of the p- region which is negatively charged, under the oxide. When a negative voltage is applied between gate and source, the channel disappears and no current can flow between the source and the drain.

If the MOSFET is a p-channel or p-MOSFET, then the source and drain are 'P+' regions and the body is an n- region. When a negative gate-source voltage is applied, it creates a channel which is positively charged at the surface of the n-region, just below the oxide.

The operation of a MOSFET can be divided into three different regions, depending upon the voltages at the terminals. The three regions of operation are cut-off, linear and saturation which are explained below [2].

- Cut-off or subthreshold mode (When $V_{GS} < V_{th}$ where V_{th} is the threshold voltage of the device and V_{GS} is the gate to source voltage)

  Under these operating conditions the transistor is turned off, and there is minimal conduction between the drain and source. However, the Boltzmann distribution of electron energies allows some energetic electrons at the source to enter the channel and flow to the drain creating a diffusion current. This subthreshold current is an exponential function of the gate to source voltage. The current between the drain and source should ideally be zero when the transistor is being used as a turned-off switch; the weak-inversion current, sometimes called subthreshold leakage is very critical in low power digital circuits [2]. However, the subthreshold current or weak inversion region is an efficient region of operation in some analogue circuits [2].

- Triode or linear region (When $V_{GS} > V_{th}$ and $V_{DS} < V_{GS} - V_{th}$)

  In the linear region where the gate to source voltage V_{GS} is greater than the threshold voltage V_{th} the transistor is turned on, and a channel has been created which allows current to flow between the drain and source. This condition is depicted in Figure 2.7. The MOSFET operates like a resistor, controlled by the gate voltage relative to both the source and drain voltages.

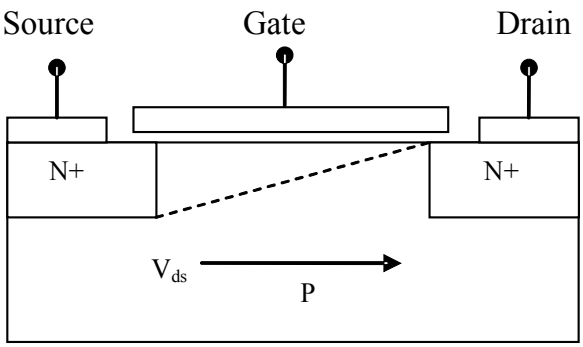


Fig. 2.7 Cross section of a MOSFET operating in the linear region.

The drain current is given by the relation [2];

$$I_D = \mu_n C_{ox} \frac{W}{L} \left((V_{GS} - V_{th}) V_{DS} - \frac{V_{DS}^2}{2} \right) \qquad \text{…….. (2.1)}$$

Where μ_n is the charge-carrier mobility, W is the gate width, L is the gate length and C_{ox} is the gate oxide capacitance per unit area.

- Saturation (When $V_{GS} > V_{th}$ and $V_{DS} > V_{GS} - V_{th}$)

In this case when the drain voltage is increased, a channel has been created, which allows current to flow between the drain and source. Since the drain voltage is higher than the gate voltage, a portion of the channel is turned off. The onset of this region is also known as pinch-off. In this region the drain current is now relatively independent of the drain voltage and the current is controlled by only the gate to source voltage. This operating condition is shown in Figure 2.8.

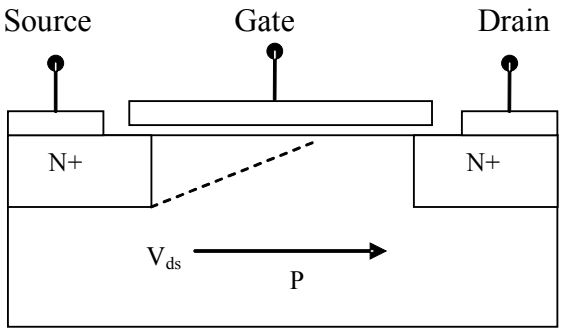


Fig. 2.8 Cross section of a MOSFET operating in saturation.

The equation for the current in this region is given by [2];

$$I_D = \frac{\mu_n C_{ox}}{2}\frac{W}{L}(V_{GS} - V_{th})^2 \qquad \text{…….. (2.2)}$$

For larger drain biases, the length of the inverted drain region decreases with increase in drain bias leading to channel length modulation which is discussed later in this chapter. Channel length modulation leads to an increase of current in the channel with drain bias and hence a lower output resistance for the MOSFET [2]. Equation 2.2 can be multiplied by $(1 + \lambda V_{DS})$ to take account of the channel length modulation effect [2]; where λ is the channel length modulation parameter.

- Body Effect

The body effect describes the changes in the threshold voltage by the change in the source-bulk voltage, approximated by the following equation;

$$V_{TN} = V_{TO} + \gamma\left(\sqrt{V_{SB} + 2\phi} - \sqrt{2\phi}\right) \qquad \text{…….. (2.3)}$$

where V_{TO} is the zero substrate bias, γ is the body effect parameter, and 2φ is the surface potential parameter [2].The body (substrate) can be operated as a second gate, and is referred to as the back gate and the body effect is sometimes called the back-gate effect [2, 3].

The Figure 2.9 shows the family of current/voltage curves of an n-MOSFET for increasing V_{DS} (drain - source voltage) for different V_{GS} (gate - source voltage). The linear region and saturation region are marked in the figure.

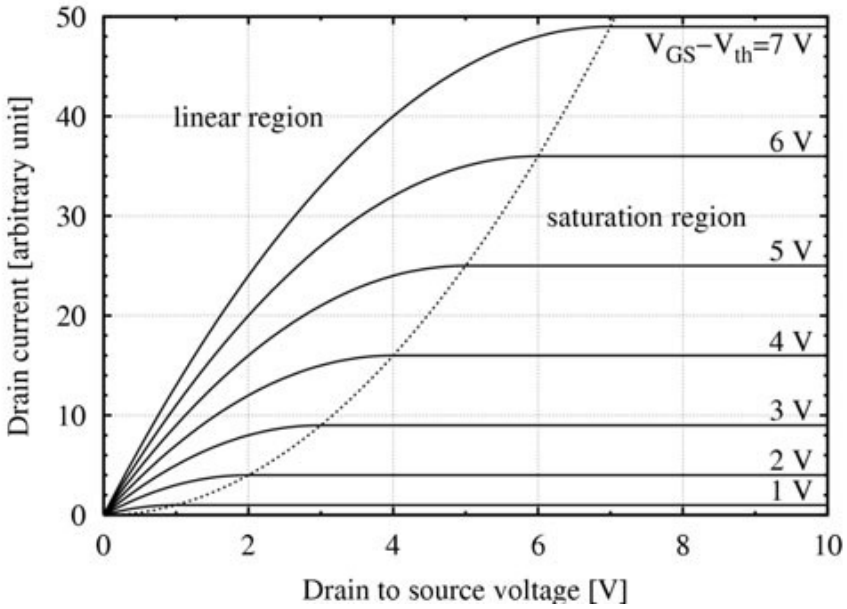


Fig. 2.9 The family of I_{DS} –V_{DS} curves for different V_{GS}.

In summary, the gate terminal of a MOS transistor controls the flow of current between source and the drain terminals. In the case of n-MOS, when the gate is '1', the transistor is ON and there is a conducting path from source to drain whereas p-MOS transistor is turned ON when the gate is at low voltage ('0').

2.1.6 CMOS

A balance of low power and high throughput are the main goals which the microelectronics industry must address in satisfying the demand for more advanced applications in the consumer market sector for portable equipment in the modern world [6]. The designer can make optimizations at all levels of the design space, which have a cumulative effect on the overall system power reduction. Much work has been concentrated on architecture, algorithm, and system-level power minimization. The technology for IC manufacturing is the only level that the designer has limited control to meet the constraints.

Complimentary Metal Oxide Semiconductors (CMOS) circuits were invented in 1963 by Frank Wanlass at Fairchild Semiconductor as a low-power alternative to Transistor Transistor Logic (TTL) [6, 7]. The first CMOS integrated

circuits were made by Radio Corporation of America (RCA) in 1968 by a group led by Albert Medwin [6]. CMOS found applications in the watch industry and in other fields where battery standby capability was more important than speed. After around twenty-five years, CMOS has become the predominant technology in digital integrated circuits and still maintains this position [8]. The main advantages of CMOS over TTL are its energy efficiency, smaller area occupation, comparable operating speed and manufacturing costs [7]. CMOS benefits from the geometric scaling of dimensions that comes with every new technology node associated with semiconductor processing. Besides all these advantages, low-power dissipation and larger integration densities, compared to bipolar junction transistors, has made it the mainstay of the microelectronics world [7].

CMOS uses a combination of p-type and n-type MOSFETs on the same substrate to implement logic gates and other digital circuits found in computers, telecommunications and signal processing equipment. Figure 2.10 shows the structure of a CMOS.

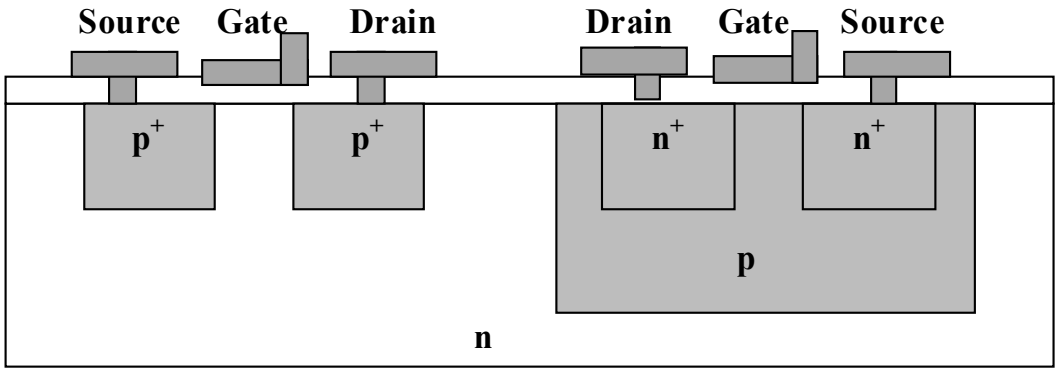


Fig. 2.10 Structure of a CMOS.

As an example of CMOS transistors used in a digital circuit application consider the inverter circuit shown in Figure 2.11, with an n-type MOSFETs in the pull-down network between the output and the lower-voltage power supply or ground and a p-type in a pull-up network between the output and the higher-voltage rail or power supply.

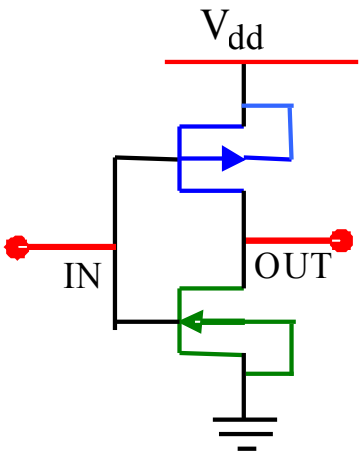


Fig. 2.11 Schematic of an inverter logic gate.

Pull-up charges the load capacitance when there is a direct path from supply voltage source (V_{dd}) to the output, and discharges (pull-down) when there is a direct path from output to ground. In CMOS since the p-type transistor network is complementary to the n-type transistor network, when the n-type is off, the p-type is on, and vice-versa. The advantage here is that in a CMOS circuit pull-up and pull-down never happen at the same time [7]. Hence the static power of CMOS based circuits are minimal compared to the equivalent circuit based on passive components. CMOS dissipates power only when switching (dynamic power).

The switching power dissipated by a CMOS circuit is given by $P = CV^2f$; where C is the capacitance, V the voltage and f, is the switching frequency [7]. Both n- and p- MOS transistors have a threshold gate-to-source voltage, below which the current through the device drops exponentially. For older technology generations CMOS circuits were operated at supply voltages around 5V which is much larger than their threshold voltages. Supply voltage scaling has been evolved as a technique to reduce the power consumption of circuits for newer applications [9-13]. However, the associated disadvantage with these techniques is increased subthreshold current and hence increased power consumption.

2.1.7 Performance Enhancement Techniques

To improve the performance of devices, a technique called scaling has been adopted by the microelectronics industry where the geometric dimensions of the transistors are scaled down [14]. Scaling, in general will improve the speed of a MOSFET along with the improvement in packing density (area). For example, shorter gate length and gate oxide thickness will improve the transconductance (transconductance is the change in the drain/source current divided by the change in the gate/source voltage with a constant drain/source voltage) which is an important performance metric of devices. Smaller MOSFETs have smaller gate areas, and thus lower gate capacitance. These factors will contribute to faster switching times, and thus higher circuit speeds and lower energy per switching event (due to load capacitance) [15]. Another reason for MOSFET scaling is reduced area, leading to reduced cost (more transistors per unit area). Smaller MOSFETs can be packed more densely, resulting in either smaller chips or chips with more computing power in the same area. Since fabrication costs for a semiconductor wafer are relatively fixed, the cost per integrated circuits is mainly related to the number of chips that can be produced per wafer. Hence, smaller ICs allow more chips per wafer, reducing the price per chip. However, there are several critical issues associated with scaling which are discussed in Section 2.1.8 below.

2.1.8 Scaling Issues

Newer generation MOSFETs are smaller than 100 nanometers in size due to the relentless scaling of devices by the microelectronics industry to keep the Moore's Law alive [16]. Manufacturing of MOSFETs with channel lengths in the order of nanometers is a challenge, and the difficulties associated with the fabrication of microelectronic devices are the only limiting factor in advancing integrated circuit technology [13]. When the dimensions become smaller, the issues related to short channel effects at process, device and circuit level are shown in Figure 2.12. In recent years, devices with channel lengths in the range of few tens of nanometers have been fabricated. However, the advantages exploited by the microelectronics industry through scaling are reaching a plateau as discussed below.

Sub- threshold Conduction

To improve the device drain current, the threshold voltage of the MOSFET has to be reduced [17]. This can be done by scaling or by adjusting other process conditions for example doping etc. To maintain reliability, the supply voltage applied to the gate must also be scaled because of the scaled transistor dimensions [17]. As the threshold voltage is reduced, the transistor cannot be completely turned off; it operates in weak-inversion mode, with a leakage current between source and drain, called subthreshold leakage current, which leads to power being consumed. Modern day high performance devices in sub-100nm technology node consume leakage power almost equal to half of the total power consumption [17].

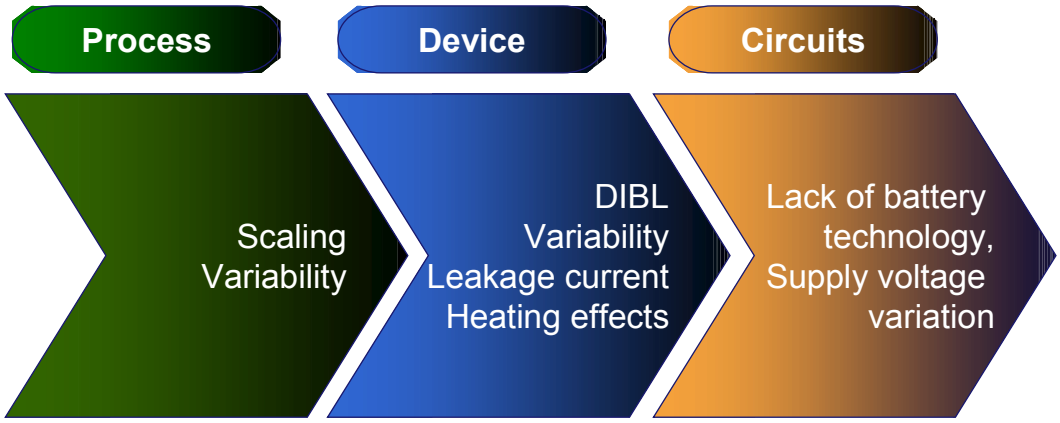


Fig. 2.12 Short Channel Issues Transmitted from Process to Device to Circuits.

Heat Production

The ever-increasing density of MOSFETs on an integrated circuit is creating problems of substantial localized heat generation that can impair circuit operation. Circuits operate more slowly at high temperatures, and have reduced reliability and shorter lifetimes [18]. Heat sinks and other cooling methods are now required for many integrated circuits including microprocessors to dissipate the generated heat. As their on-state resistance rises with temperature, the power loss on the junction rises correspondingly, generating further heat. When the heat sink is not able to keep the temperature low enough, the junction temperature may quickly and uncontrollably rise, resulting in thermal runway of the device [18].

Gate Oxide Leakage

When the transistor is scaled to improve the performance, the gate oxide thickness also has to be scaled and made as thin as possible to increase the channel transconductance and performance when the transistor is turned 'ON'. It also helps to reduce subthreshold leakage when the transistor is turned 'OFF' However, with current gate oxides having a thickness of around 1.2 nm [2] the quantum mechanical phenomenon of electron tunneling occurs between the gate and channel, leading to increased power consumption .

Oxide Breakdown

As the gate-oxide is scaled down to improve the performance of transistors as discussed in previous section, breakdown of the oxide and oxide reliability becomes more of a concern. Higher fields in the oxide increase the tunnelling of carriers from the channel into the oxide [19]. These carriers slowly degrade the quality of the oxide and lead to its failure. This effect is referred to as time dependent dielectric breakdown (TDDB) [20].

Channel Length Modulation

As the drain voltage is increased, the depletion layer width in the channel increases and hence the carrier channel will be shortened leading to an increased drain current. This phenomenon is called the channel length modulation [1]. The channel length modulation effect increases in short channel devices with low-doped substrates. With increasing drain voltage there is an extreme case that the channel length reduces to zero and this effect is called punch through [2].

Punch Through

Punch through in a MOSFET is an extreme case of channel length modulation where the depletion layers around the drain and source regions merge into a single depletion region. The drain current hence becomes a function of drain source voltage. Punch through causes a spontaneously increasing current with increasing drain-source voltage. This effect is undesirable as it increases the output conductance and limits the maximum operating voltage of the device [2]. This effect

could be controlled by using shallow junction extensions. In addition, halo doping can be used which is a very thin heavily doped regions of the same doping type as the substrate near the junction walls to limit the extension of depletion regions [21].

Drain Induced Barrier Lowering (DIBL)

When the output conductance and threshold voltage depends on drain voltage the effect is called DIBL. This effect occurs in devices where only the gate length is reduced without properly scaling the other dimensions. It is observed as a variation of the measured threshold voltage with reduced gate length. The threshold variation is caused by the increased current with increased drain voltage as the applied drain voltage controls the inversion layer charge at the drain, thereby competing with the gate voltage. This effect is due to the two-dimensional field distribution at the drain end and can typically be eliminated by properly scaling the drain and source depths while increasing the substrate doping density [2, 21].

Avalanche Breakdown and Bipolar Parasitic Action

As the electric field in the channel is increased, avalanche breakdown occurs in the channel at the drain. This avalanche breakdown increases the current. In addition, there is a parasitic bipolar action taking place. Holes generated by the avalanche breakdown move from the drain to source underneath the inversion layer through the substrate. The hole current forward biases the source-bulk p-n diode (n-MOSFET) so that electrons are injected as minority carriers into the p-type substrate underneath the inversion layer. These electrons arrive at the drain and again create more electron-hole pairs through a process called avalanche multiplication. The positive feedback between the avalanche breakdown and the parasitic bipolar action results in breakdown at lower drain voltage [22].

Velocity Saturation

As devices are scaled to improve the performance, the electric field in the channel increases and the carriers in the channel travel with an increased velocity. However at high fields the relation is non linear reaching a saturation velocity. This effect is due to scattering of highly energetic electrons due to optical phonon

emission. The scattering effectively reduces the mean free time of carriers travelling in the channel [2].

Process Variations

Most modern digital VLSI designs operate with GHz frequencies. Hence precision in the design and manufacturing process is very critical. During chip manufacturing, random process variation can affect the size of the transistor, which becomes a greater percentage of the overall transistor size as the dimension shrink. With MOSFETS becoming smaller, the number of atoms in the semiconductor material that produce many of the transistor properties is becoming fewer thereby amplifying the dependence of the transistor characteristics with process parameters. The transistor characteristics become less deterministic. This statistical variation increases design difficulty [23]. Design corners based on circuit applications can be identified once the distribution of the variation of parameters is known. The variation of the threshold voltage for technology generations ranging from 250nm down to 45nm are shown in Table 2.1.

L (nm)	Nominal Vth (mV)	% Change from the nominal
250	450	4.7
180	400	5.8
130	330	8.2
60	300	9.3
65	280	10.7
45	200	16

Table. 2.1 Higher relative variability of threshold voltage for newer technologies [24].

The unnecessary parameter fluctuations can be differentiated based on the variability sources which can be physical or environmental. The physical source can be the manufacturing process, manufacturing equipment, electron migration,

changes in the characteristics of devices and/or interconnect lines due to ageing. The environmental sources can be due to the change in supply voltage, temperature (operating conditions), local coupling, peculiar design implementation etc. The process variation which is a physical variation can occur from wafer to wafer or lot to lot (both are inter process variations). It can also be between die to die (intra process variations) [24, 25].

Wafer to wafer or lot to lot variations are global variations which can be due to materials and gas flow (linear variation) or thermal or the wafer spin process (radial variation) [24]. This type of variation mainly affects the dynamic performance of the digital systems for example switching speed, dynamic power and gain in analog systems. There can be reticle variation due to the optical processes. Other sources of variations include position and proximity variations arising from the layout. These variations can be extracted and modeled and since they are systematic they are called predictable variations. Variations arising from layout can be optimized by several methods, for example mask compensation [24].

Local variations between identically laid-out devices arise from random microscopic processes variations. These variations are randomly distributed and are unpredictable. For 130nm technology, local variations constitute around 30% of the overall variations [24]. The random distributions arise from the variation of process parameters for example impurity concentration densities, oxide thickness and diffusion depths which result from varying operating or environmental conditions during the deposition or diffusion of the impurity dopants. The fluctuations in the process parameters may result in the variation of sheet resistance and threshold voltage. However, the variation of threshold voltage can also be due to the variation of surface charge or change in the oxide thickness. The varying resolution of the photolithographic process may lead to the variation of geometric dimensions which may result in the variation of transistor output characteristics for example, drain current. These variations are not correlated because the width of the transistor (W) is determined in the field oxide step while the effective length of the channel (L_{eff}) is defined in the poly and source-drain diffusion steps [24].

The variations will impact the performance of circuit, which may exhibit wider variability leading to the degradation of yield in modern technologies and

applications. The circuit parametric variations for a 130nm technology process are shown in the Table 2.2. For a given operating temperature L_{eff}, V_{th} and V_{dd} are the most dominant variation sources of a logic gate. V_{dd} and temperature are run-time variation sources [24]. Figure 2.13 shows the impact of process variations on the delay of a 4 bit adder fabricated using 130nm technology [24].

Process	Change	Interconnect	% change
L_{eff}	16.70%	ε	3
V_{th} (V)	30%	ρ	30
t_{ox} (° A)	10%	*w*	20
R_{ds} (Ω)	10%	s (nm)	20
Run-time		t (nm)	10
V_{dd} (V)	10%	*h*	10
T (°C)	25 -100	R_{via} (Ω)	20

Table. 2.2 Sources of major variation at 130nm technology node ($3\sigma/\mu$) [24].

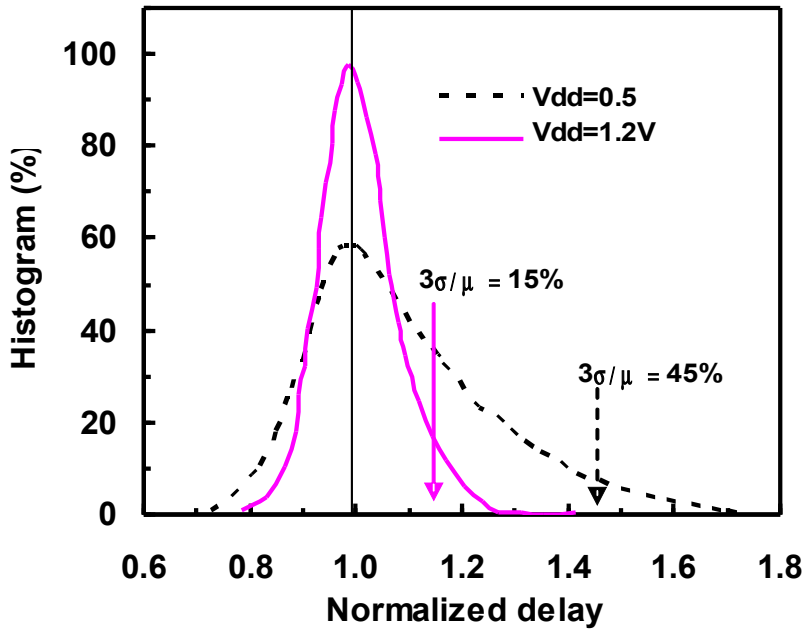


Fig. 2.13 Monte Carlo simulations of a 4-bit adder – Impact of variations on delay [24].

Using the variation values shown in Table 2.2 it can be seen from Figure 2.13 that the performance variability $3\sigma/\mu$ is as large as 15% at a supply voltage of 1.2V. However, when the supply voltage is scaled to 0.5V in an aim to reduce the power consumption, the performance variability increases to 45% which shows that circuit yield degrades with voltage scaling.

The impact of variations on leakage current for a 4KB SDRAM realized with 130nm technology generation is shown in Figure 2.14. It can be observed from the figure that in comparison to the leakage current values simulated without considering channel length and threshold voltage variations, the leakage current obtained by measurement, rises exponentially at large values of supply voltages which may lead to the malfunctioning of the SRAM. This dramatic increase of leakage current in the measurement data shown in Figure 2.14 is caused by the transistors with shorter channel length and lower threshold voltage [24].

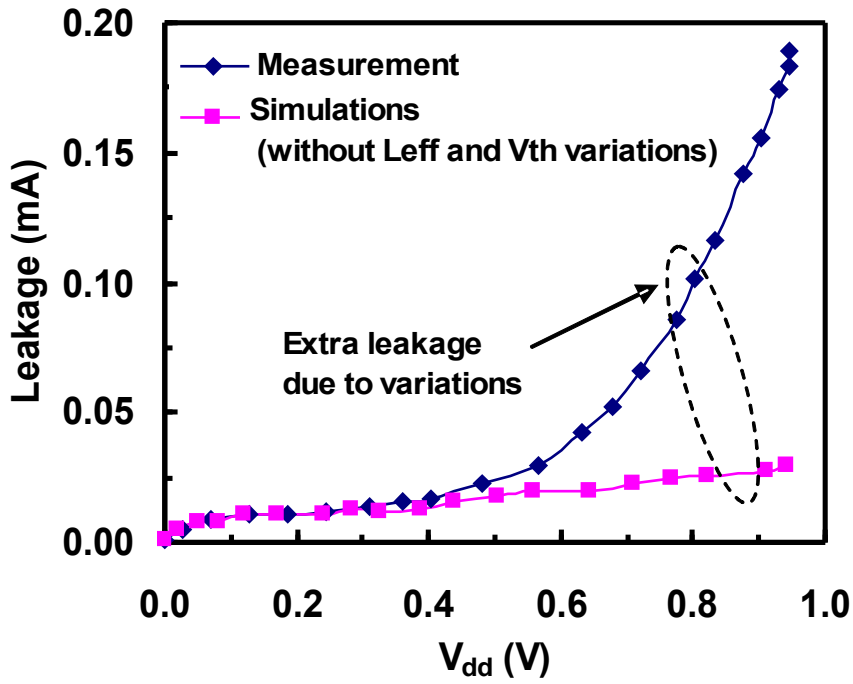


Fig. 2.14 Leakage measured from a 4KB SDRAM- Impact of variations on leakage power consumption [24].

Given that compensating for random variations is difficult, containing the effects of the random parameter variations can be considered as the most significant challenge the microelectronics industry will face in the coming technology

generations. Hence it is important to have an organized and collective effort of technology and design engineers to mitigate the effect of variations on modern day circuits. Designers need EDA tools to accurately model different sources of variations to guarantee improved yield which could be called as 'Process Aware Design' (PAD). PAD could exploit the issues of variability to achieve trade-offs between power and circuit throughput thereby saving the cost from using the worst-case design approach [26]. Hence the first step is to study a given variation and its impact on circuit performance. However the amount of variation expected from various technologies for example strained-Si, SOI, SGOI etc are different.

There are other issues with scaling which affect the IC as a whole, for example interconnect capacitance. The switching time of long channel CMOS circuits was roughly proportional to the gate capacitance. However, with transistor scaling and high packing density on the chip, interconnect capacitance (the capacitance of the wires connecting different parts of the chip) is contributing a high percentage to the overall capacitance. Signals have to travel through the interconnect wires, which leads to increased delay and hence lower throughput [27].

2.2 Need for new Technologies

The power delay product is a measure of the semiconductor device performance; power could be compromised for speed of operation [11]. Lowering the power consumption is as vital a requirement as it is to increase the speed in future circuits [11]. The conventional method to improve performance is to scale down the geometric dimensions of devices to increase the saturation current [17]. However, the leakage power approaches the switching power when transistor dimensions approach the nanometer range. As a result, these devices dissipate considerable power even when not switching. Power reduction using new materials and system design techniques is critical to sustain Moore's Law [16] and saving the transistors from severe short channel effects. Sections 2.1.7 and 2.1.8 described the conventional scaling technique and its associated issues. The following section discusses the issues associated with the conventional method used to improve performance and about the new materials, structures and technologies which the

industry is contemplating to combat the increasing leakage power and other scaling issues.

2.3 New Technologies

In order to bypass the issues associated with device scaling as a means to improve device performance, industry has been considering the use of new material and device structures, Developments in these areas are outlined in the following sections.

2.3.1 Materials

To tackle the undesired electrical and geometrical effects of deep sub-micron transistors on new circuit applications, together with the use of new circuit techniques, new materials and structures which could give high carrier mobility and which can be operated at lower supply voltage have been adopted. However, a sharp change from Si is not likely to be accepted by the industrial community as Si enjoys some inherent advantages such as low cost, ease of manufacturing etc [29]. Some of the new technologies, materials and structures which are under development are included in ITRS [28]. The new technologies which are complementary to device scaling and pertinent to modern electronic applications are discussed below.

2.3.1.1 Si on Insulator (SOI)

As the name suggests the SOI uses Si on Insulator (an oxide usually) technology instead of conventional Si substrate for manufacturing transistors. The choice of insulator depends on the transistor application. Sapphire is used sometimes as an insulator for radiation-sensitive and military applications. For commercial applications Si oxide (SiO_2) insulator is preferred for improved performance, reduced short channel effects as it has fewer extended defects compared to other oxides. SOI can be of two types; partially depleted and fully depleted, the difference being the thickness of the top layer Si. The thickness of the insulating layer and topmost Si layer also varies widely depending on the applications and operating environment.

Compared to conventional bulk Si SOI is superior in several aspects. The advantages of SOI which are reported [30] include;

- Lower parasitic capacitance due to the isolation from the bulk Si.
- Power consumption at matched performance.
- Resistance to latch up due to complete isolation of the n- and p- well structures.
- No retooling of fabrication facility required.
- SOI substrates compatible with Si fabrication processes.

Advantages from the manufacturing side include no special machines or equipment needed for SOI compared to conventional Si which reduces the cost considerably and makes it a very interesting technology for microelectronics. However, there is a drastic increase in substrate cost. Furthermore, there are still concerns with the stress in the Si layer above the insulator and due to the lower thermal conductivity of the oxide underneath the Si layer which has comparatively higher thermal conductivity [31]. SOI devices can be manufactured using several different methods for example Separation by IMplantation of Oxygen (SIMOX) [32], Wafer Bonding [33], ELTRAN [33] , Seed methods [21] etc.

2.3.1.2 SSOI/ SGOI (Strained Si on Insulator/ Si Germanium on Insulator)

In these technologies the advantages obtained from straining the channel of the transistor are combined with the benefits obtained from SOI [34]. The strained Si (strained-Si) technology is explained, later, in detail in the Section 2.3.1.5.

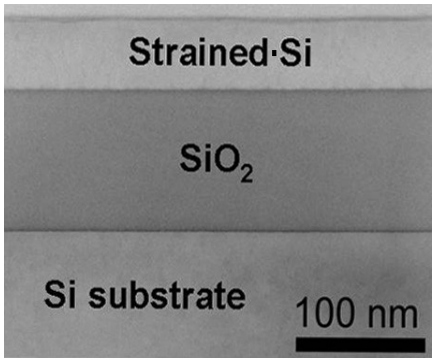


Fig. 2.15 TEM image of an SSOI MOSFET [34].

Relaxed $Si_{1-x}Ge_x$-on-insulator (SGOI) is a very promising technology as it combines the benefits of two advanced technologies. Along with the advantage the SOI devices offer, as explained in the previous section, the additional advantages obtained from SiGe based devices make them suitable for a wide range of applications such as mobile communications. The same methodology can be used to combine high mobility strained-Si or strained-$Si_{1-x}Ge_x$ on SGOI substrates. Figure 2.15 shows the cross section of an SGOI MOSFET using Transmission Electron Microscope (TEM) [34, 35].

2.3.1.3 Silicon Carbide (SiC)

Silicon Carbide (SiC) is a man-made ceramic compound of Si and carbon which can be used as a semiconductor. It finds a wide range of applications which include automotive parts, ultraviolet detectors, jewels, weapons, furnaces etc [36]. The semiconductor devices based on SiC exhibit high resistance until the voltage across them reaches a threshold voltage, after which the resistance drops to a very low value and maintains this until the applied voltage drops below the threshold again. Hence the transistors based on SiC are used in applications where high power is necessary [36].

SiC carbide is used for blue LEDs, ultra-fast high-voltage Schottky diodes, MOSFETs and high temperature thyristors for high power switching [37]. Due to its wide band gap, SiC-based parts are capable of operating at high temperature (over 350 °C), which together with good thermal conductivity of SiC reduces problems of cooling high power devices.

2.3.1.4 SIT and SOS

Transistors have been developed using other methods and materials for example, Static Induction Transistor (SIT) which is a short-channel FET used in High Power RF systems [38]. Si-On-Sapphire, which is similar to SOI but with sapphire instead of oxide as an insulator is another technology which is not very common in the microelectronics world [39].

2.3.1.5 Strained silicon (Strained-Si)

In recent years strained-Si devices have become the main stream of microelectronics industry as a potential candidate to complement device scaling due to its inherent advantages in comparison to Si devices and its compatibility with the existing Si process technology [40]. Due to its high potential, ITRS [28] has already included it in their update on the semiconductor roadmap. Several companies for example IBM, Intel, AMD and Amberwave have demonstrated the advantages of strained-Si over conventional technologies in terms of saturation current and threshold voltage of devices. There are several reports which demonstrate the improvements of n-MOS and p-MOS drive currents due to the improvement in electron and hole mobilities respectively. IBM demonstrated an improvement of 70% in electron flow which makes the devices faster by around 35% for identically layout devices (without shrinking its dimensions) [41]. The performance enhancements obtained by straining the devices benefit the industry by enabling circuits with a high performance capability to be manufactured which are suitable for applications such as gaming graphics cards, GPS navigation and mobile communication applications. Strained-Si, compared to previously mentioned technologies such as SOI, enjoys inherent advantages for example compatibility to existing Si technology and lesser heating effects which are very critical in SOI devices where bulk heat sinks have to be used [42]. Given that the microelectronics industry is much interested in strained-Si due to its inherent advantages, this thesis concentrates on circuits based on strained-Si technology and hence it is discussed in detail in Chapter 3.

2.3.1.6 High Dielectric Constant Materials (High-K)

The new materials discussed in the previous section which improves the performance of devices apply to the substrate part of the MOSFET to combat scaling. According to the generic rule for scaling, to improve the performance of transistors, when transistors are scaled, the thickness of the silicon dioxide (SiO_2) gate dielectric has been decreased to increase the gate capacitance thereby improving the drive current and device performance as a whole [43]. However, as the oxide layer thickness is reduced below 2nm ‘tunneling effects’ become an issue [2]. The effect of tunneling is to increase the leakage currents drastically, leading to

very high power consumption and reduced device reliability. According to ITRS [28] high performance devices on a 90nm technology node require an oxide thickness of 1.2nm leading to high leakage current through tunneling. To combat these effects there is a need for different gate oxide material than SiO_2 which has been used for decades.

The implementation of high-k gate dielectrics is one of several strategies developed to allow further miniaturization of integrated circuits to follow ITRS proposed specifications [12]. The 2006 ITRS roadmap predicts the implementation of high-k materials to be commonplace in the industry by 2010. Replacing the SiO_2 gate dielectric with a high-k material allows increased gate capacitance without any critical leakage effects [44]. High-k dielectric refers to a material with a high dielectric constant (k) compared to SiO_2.The dielectric between the gate electrode and the channel of a MOSFET can be modeled as a parallel plate capacitor [2].

Ignoring quantum mechanical and depletion effects from the Si substrate and gate, the capacitance C of this parallel plate capacitor is given by [2];

$$C = \frac{k\varepsilon_0 A}{t} \qquad \text{…….. (2.4)}$$

Where, A is the capacitor area,

k is the relative dielectric constant of the material (3.9 for SiO_2),

ε_0 is the permittivity of free space,

t is the thickness of the capacitor oxide insulator.

From Equation 2.4, it can be seen that the capacitance can be increased by either increasing k or decreasing t. Historically, t was decreasing which leads to undesired effects for example, current leakage. Hence the alternative to improve the capacitance is to increase k by replacing SiO_2 with a high-k material. In this case a

thicker gate layer could be used which can reduce the leakage current flowing through the structure as well as improving the gate dielectric reliability.

A typical gate dielectric structure with SiO_2 is compared with a high-k dielectric structure as shown in Figure 2.16.

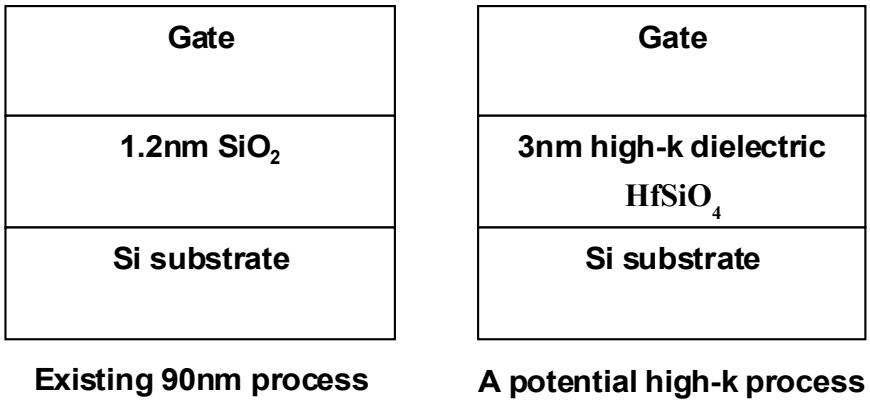


Fig. 2.16 SiO_2 gate dielectric structure compared to a high-k dielectric structure [45].

The drive current, Id for a high-k MOSFET can be approximated by Equation 2.5 [2].

$$Id = \frac{W\mu C_{inv}}{L}\left(V_G - V_T - \frac{V_D}{2}\right)V_D \qquad \text{…….. (2.5)}$$

Where, W is the width of the transistor channel,

L is the channel length,

μ is the channel carrier mobility (assumed constant here),

C_{inv} is the capacitance per unit area associated with the gate dielectric when the underlying channel is in the inverted state,

V_G is the voltage applied to the transistor gate,

V_D is the voltage applied to the transistor drain,

V_T is the threshold voltage.

It can be seen from the Equation 2.5 that the reduction in the channel length or an increase in the gate capacitance will result in an increased saturation current.

When new materials are being sought to replace the conventional oxide dielectric, the material quality, complexity and its seamless integration into the existing manufacturing process are the greatest challenges. Other long time challenges include, obtaining the same stability, reliability and breakdown voltage as SiO_2. An important consideration is the barrier height of the new gate oxide; the difference in conduction band energy between the semiconductor and the oxide (and the corresponding difference in valence band energy) will also affect the leakage current level.

2.3.2 Structures

Scaling has been the driver for high packing density chips with performance as bonus. Apart from introducing new materials for the manufacture of transistors to improve the functionality per unit area, new transistor structures are also finding importance in the microelectronics industry to keep the Scaling Law alive [16]; several of these new structures are discussed below.

A multi-gate device or Multi-gate Field Effect Transistor (MuGFET) refers to a MOSFET which incorporates more than one gate into a single device. The multiple gates may be controlled by a single gate electrode, wherein the multiple gate surfaces act electrically a single gate, or by independent gate electrodes. A multi-gate device employing independent gate electrodes is called a Multiple Independent Gate Field Effect Transistor or MIGFET [46]. The primary roadblock to its widespread implementation is manufacturability, as both planar and non-planar designs present significant challenges, especially with respect to lithography and patterning [46].

In a multi-gate device, the channel is surrounded by several gates on multiple surfaces, allowing more effective suppression of the "OFF" state leakage current. Multiple gates also allow enhanced current in the "ON" state, also known as the drive current. These advantages translate into lower power consumption and enhanced device performance. Non-planar devices are also more compact than

conventional planar transistors, enabling higher transistor density which translates into smaller chip sizes.

However, the primary challenges to integrating non-planar multi-gate devices into conventional semiconductor manufacturing processes include [46]:

- Fabrication of thin Si “fins" tens of nanometers wide.
- Fabrication of matched gates on multiple sides of the fin.

There are different types of Multi Gate transistors which can be classified based on the architecture (planar vs. non-planar design) and number of channels/gates (2, 3, or 4). Several of these variations will be discussed below.

2.3.2.1 Planar Double Gate Transistors

Planar double-gate transistors employ conventional planar (layer by layer) manufacturing processes to create double-gate devices, avoiding more stringent lithography requirements associated with non-planar vertical transistor structures. In planar double-gate transistors the channel is sandwiched between two independently fabricated gate/gate oxide stacks. The primary challenge in fabricating such structures is achieving satisfactory self-alignment between the upper and lower gates [46].

2.3.2.2 FinFETs

The FinFET is based on a single gate transistor design and is a non-planar double-gate transistor built on an SOI substrate. The distinguishing characteristic of the FinFET is that the conducting channel is wrapped around a thin Si "fin", which forms the body of the device. The dimensions of the fin determine the effective channel length of the device [47]. A 25-nm transistor which operates on 0.7 Volt was demonstrated in December 2002 by Taiwan Semiconductor Manufacturing Company. The structure of FinFET is shown in Figure 2.17.

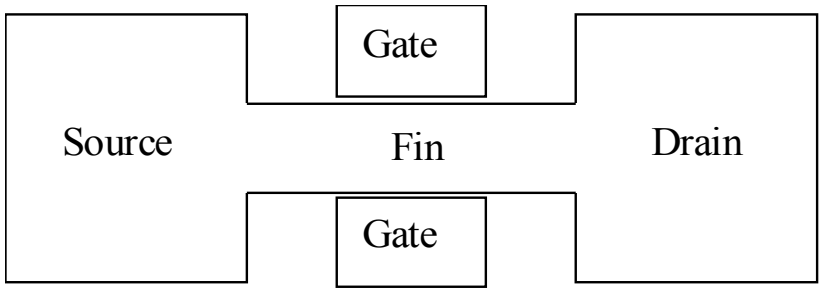


Fig. 2.17 Structure of a double-gate FinFET device [45].

2.3.2.4. Gate-all-around (GAA) Field Effect Transistors

Gate-all-around FETs are similar in concept to FinFETs except that the gate material surrounds the channel region on all sides. Depending on the design, gate-all-around FETs can have two or four effective gates [45].

2.4 Design Methods for High Performance and Low Power

Apart from the technology initiatives using new materials and structures which are discussed in the previous sections to improve the performance of transistors thereby improving overall system performance; different design methodologies are used by the designers at a higher abstraction level to achieve high throughput and low power depending on circuit applications. Figure 2.18 shows several design and technology methods which are used by technologists and designers to achieve high performance systems. The technology methods shown in the figure, that is, strain and scaling have been discussed in previous sections. The design methods namely, Globally Asynchronous Locally Synchronous (GALS), dual V_{th} designs, stacking, are discussed below.

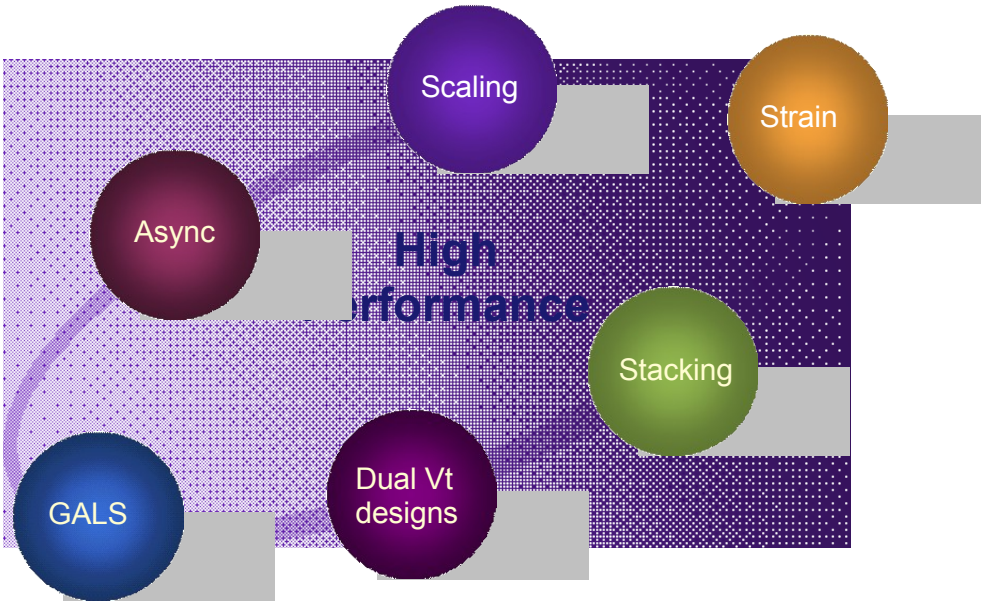


Fig. 2.18 Technology and Design methods to achieve High Performance.

2.4.1 Asynchronous techniques

The peculiarity of asynchronous circuit unlike synchronous circuits is that it is not governed by a clock circuit or global clock signal. It needs only to wait for the signals which indicate the completion of instructions and operations. These signals are specified by simple data transfer protocols [48]. Some of the disadvantages of technology scaling such as increased power consumption [12] can be removed using asynchronous techniques. Globally Asynchronous Locally Synchronous (GALS) [50] techniques are gaining importance where applications are neither synchronous nor asynchronous for example, ATM networks, automobile and aircraft controllers etc. The advantages offered by asynchronous technique include:

- Lower power consumption due to the fact that only the computational active parts of the circuit dissipates power. However, sometimes, asynchronous circuits may require more area, which can result in increased power consumption if the underlying process has poor leakage properties (for example, DSM processes used prior to the introduction of high-K dielectrics) [48].

- Circuit speed is adapted to changing temperature and voltage conditions rather than being locked at the speed mandated by worst-case assumptions [48].
- Immunity to transistor-to-transistor variability in the manufacturing process, which is one of the most serious problems facing the semiconductor industry as die sizes shrink [49].
- Generate less electromagnetic interference. Synchronous circuits create a great deal of EMI in the frequency band at (or very near) their clock frequency and its harmonics; asynchronous circuits generate EMI patterns which are much more evenly spread across the spectrum [48].

2.4.2 Dual V_{th} designs

To reduce leakage power consumption, designers have introduced the concept of dual V_{th}. This approach reduces leakage power without sacrificing the timing performance of the design by using cells with increased switching voltage thresholds to replace as many of the regular threshold cells as possible without degrading timing [50]. The high V_{th} cells are identical to the low V_{th} cells except for processing changes. To maintain performance, all gates on the critical path are assigned low V_{th} and the gates on the non-critical paths are assigned high V_{th}. The advantage is that dual V_{th} designs can reduce leakage power and withstand process variability to some extent [50] in both standby mode and active mode of the circuit.

2.4.3 Stacking

The stacking of transistors in a circuit is an effective technique to reduce the subthreshold leakage current (static power dissipation) in its standby-mode [51]. The subthreshold current flowing through the stack of transistors decreases with an increase in the number of the `OFF` transistors. However the disadvantage is that the speed will be reduced. Figure 2.19 shows the schematic of a normal and a stacked inverter where the transistors are stacked in the pull-down network. However, the stacking can be done in pull-up network as well; the method of stacking is discussed in detail in Chapter 4.

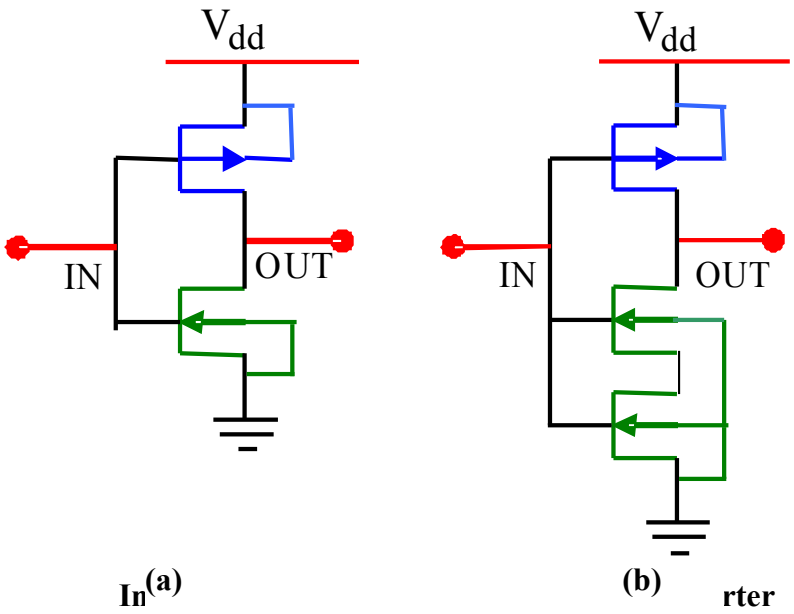


Fig. 2.19 Schematic of an inverter (a) and a stacked inverter (b).

2.5 Summary

This chapter briefly reviews semiconductor materials and device physics. The techniques to improve device performance were also discussed, in particular the use of scaling which is not only the predominant mechanism to improve device performance but also to increase the functionality per unit area, thereby reducing manufacturing costs. However, there are numerous issues associated with scaling, and subsequently the novel methods being investigated by industry to alleviate the problems through the use of new materials and device structures were also described. Some of the new technologies described include strained-Si, SOI, SSOI, SGOI, FinFETs etc. The advantages and disadvantages related to each technology were discussed. Variability is seen to be a generic challenge to all deep sub-micron devices which reduces the yield of circuits.

The final section in the chapter reviewed some of the current design techniques used to improve overall system performance and also reduce power dissipation. However, in order to extract better performance from a given technology node it has been deemed necessary to optimise circuit design around specific details of the new technologies, considering device geometries, drive current, substrate leakage, speed, etc.

References

[1] D. Long, *Energy Bands in Semiconductors*. New York: Interscience Publishers, 1968.

[2] S. M. Sze, *Physics of Semiconductor Devices*. New York: Wiley, 1981.

[3] B. G. Streetman, *Solid State Electronic Devices*. Englewood Cliffs, N.J.: Prentice-Hall, 1980.

[4] R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*. New York: Wiley, 1977.

[5] K. R. Botkar, *Integrated Circuits*. Newdelhi: Khanna, 1991.

[6] D. Lancaster, *CMOS Cookbook*. Indianapolis: H.W. Sams, 1977.

[7] N. H. E. Weste, D. Harris, and A. Banerjee, *Principles of CMOS VLSI Design: A Circuit and Systems Perspective*. New Delhi: Pearson Education, 2007.

[8] M. R. Lin, "Driving CMOS technology toward 10nm transistor gate length" Thesis (Ph. D.), Massachusetts Institute of Technology, Microsystems Technology Laboratories, 2002.

[9] P. R. Meer, A. Staveren, and A. H. M. Roermund, *Low-Power Deep Sub-micron CMOS Logic: Subthreshold Current Reduction*. London: Kluwer Academic, 2004.

[10] S. G. Narendra and A. P. Chandrakasan, *Leakage in Nanometer CMOS Technologies*. New York: Springer, 2006.

[11] J. B. Kuo, J.-H. Luo, and I. NetLibrary, *Low-voltage CMOS VLSI circuits*. New York: Wiley, 1999.

[12] R. Green, "Emerging design and technology challenges for sub-100 CMOS transistor scaling", Thesis (Ph. D.), Massachusetts Institute of Technology, Microsystems Technology Laboratories, 1998.

[13] R. A. Haken, *Submicron CMOS Technology Scaling Issues*. Stanford: University Video Communications, 1989.

[14] P. M. Zeitzoff, "Transistor scaling progresses, but new challenges loom", *Solid-State Technology*, vol. 48, pp. 30-33, 2005.

[15] M. Elgebaly, “Energy-efficient Design for Deep Sub-micron CMOS VLSIs”, Thesis (Ph. D.), University of Waterloo, 2005.

[16] G. E. Moore, "Cramming more components onto integrated circuits", *Proceedings of the IEEE*, vol. 86, pp. 82-85, 1998.

[17] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-Power CMOS Digital Design", *IEEE Journal of Solid-State Circuits*, vol. 27, pp. 473-484, 1992.

[18] L. L. Liou, B. Bayraktaroglu, and C. I. Huang, "Theoretical thermal runaway analysis of heterojunction bipolar transistors: junction temperature rise threshold", *Solid-State Electronics*, vol. 39, pp. 165, 1996.

[19] D. K. Roy, *Tunnelling and Negative resistance Phenomena in Semiconductors*. New York: Pergamon Press, 1977.

[20] J. R. Lloyd, E. Liniger, and T. M. Shaw, "Simple model for time-dependent dielectric breakdown in inter- and intralevel low-k dielectrics", *Journal of Applied Physics*, vol. 98, pp. 084109, 2005.

[21] K. R. Botkar, *Integrated circuits: A Textbook for Engineering Students*. Delhi: Khanna, 1993.

[22] R. C. Jaeger, *Microelectronic circuit design*. New York: McGraw-Hill, 1997.

[23] J. C. Levesque, “Analysis of variability in the semiconductor industry”, Thesis (Ph. D.), Massachusetts Institute of Technology, Engineering Systems Division, 2004.

[24] B. P. Wong, *Nano-CMOS Circuit and Physical Design*. New Jersey: John Wiley, 2005.

[25] A. Asenov, "Semiconductor variability requires change", *Euroasia Semiconductor*, pp. 29-34, 2007.

[26] M. Alisjahbana, “Process monitoring: predicting shifts in circuit performance based on the variability of MOSFET I (V) curves”, Thesis (Ph. D.), University of Maryland, 2000.

[27] Y. Liu, “Impact of Interconnect Variations on the Clock Skew of a Gigahertz Microprocessor”, Thesis (Ph. D.), Pittsburgh: Center for Electronic Design Automation, Carnegie Mellon University, 1999.

[28] International Technology Roadmap for Semiconductors, “Process Integration, Devices and Structures”, 2007 Ed.

[29] C. K. Maiti, N. B. Chakrabarti, and S. K. Ray, *Strained Silicon Heterostructures: Materials and Devices*. London: Institution of Electrical Engineers, 2001.

[30] G. K. Celler and S. Cristoloveanu, "Frontiers of silicon-on-insulator", *Journal of Applied Physics*, vol. 93, pp. 4955-4978, 2003.

[31] J. Olsson, "Self-heating effects in SOI bipolar transistors", *Microelectronic engineering*, vol. 56, pp. 339, 2001.

[32] J. Chen, X. Wang, M. Chen, Z. Zheng, and Y. Yu, "Formation of silicon on insulator using separation by implantation of oxygen with water plasma", *Applied Physics Letters*, vol. 78, pp. 73-75, 2001.

[33] K. Yamagata and T. Yonehara, "Bonding, Splitting and Thinning by Porous Si in ELTRAN; SOI-Epi Wafer&trade", *Materials Research Society symposia proceedings*, vol. 681, pp. 152-161, 2001.

[34] G. Taraschi, A. J. Pitera, L. M. McGill, Z.-Y. Cheng, et.al., "Strained-Si-on-Insulator (SSOI) and SiGe-on-Insulator (SGOI): Fabrication Obstacles and Solutions", *Materials Research Society symposium*, vol. 745, pp. 105-110, 2003.

[35] Z. Cheng, G. Taraschi, M. T. Currie, C. W. Leitz, et al., "LETTERS - Relaxed Silicon-Germanium on Insulator Substrate by Layer Transfer", *Journal of Electronic Materials*, vol. 30, pp. L37, 2001.

[36] Z. C. Feng and J. H. Zhao, *Silicon carbide: Materials, Processing, and Devices*. New York: Taylor & Francis, 2004.

[37] S. E. Saddow, A. Agarwal, and I. NetLibrary, *Advances in silicon carbide processing and applications*. Boston: Artech House, 2004.

[38] Y. Tang, S. Liu, S. Li, and P. Chang, "Study on combined channel parameters of a static induction transistor", *Semiconductor Science and Technology,* vol. 22, pp. 806-810, 2007.

[39] G. Marcus, K. Strohben, S. Jaskulek, A. Andreou, et al., "A monolithic isolation amplifier in silicon-on-insulator CMOS: Testing and applications", *Analog Integrated Circuits and Signal Processing*, vol. 49, pp. 63-70, 2006.

[40] S. Chattopadhyay, L. D. Driscoll, K. S. K. Kwa, S. H. Olsen, and A. G. O'Neill, "Strained Si MOSFETs on relaxed SiGe platforms: performance and challenges", *Solid-State Electronics*, vol. 48, pp. 1407-1416, 2004.

[41] P. M. Mooney, "Materials for Strained Silicon Devices", *IBM reserach report- Material Science*, pp. 1-11, 2002.

[42] K. A. Jenkins and K. Rim, "Measurement of the Effect of Self-Heating in Strained-Silicon MOSFETs", *IEEE Electron Device Letters*, vol. 23, pp. 360-362, 2002.

[43] S. Thompson, "MOS scaling transistor challenges for the 21st century", Thesis (Ph. D.), MIT Microsystems Technology Laboratories, 1998.

[44] G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High-k gate dielectrics: Current status and materials properties considerations", *Journal of Applied Physics*, vol. 89, pp. 5243, 2001.

[45] W. J. Dally and J. W. Poulton, *Digital Systems Engineering.* Cambridge: Cambridge University Press, 2000.

[46] J.-P. Colinge, *FinFETs and other Multi-gate Transistors*. New York: Springer, 2008.

[47] W. Xiong, C. R. Cleavelin, R. Wise, S. Yu, et al., "Full/partial depletion effects in FinFETs", *IEE Electronic Letters*, vol. 41, pp. 504-505, 2005.

[48] J. SparsÃ¸ and S. B. Furber, *Principles of Asynchronous Circuit Design: A Systems Perspective*. Boston: Kluwer Academic Publishers, 2001.

[49] A. Iyer and D. Marculescu, "Power and Performance Evaluation of Globally Asynchronous Locally Synchronous Processors", *Computer Architecture News*, vol. 30, pp. 158-170, 2002.

[50] A. Agarwal, C. H. Kim, S. Mukhopadhyay, and K. Roy, "Leakage in Nano-Scale Technologies: Mechanisms, Impact and Design Considerations", *Design Automation Conference*, pp. 6-11, 2004.

[51] Z. Chen, M. Johnson, L. Wei, and K. Roy, "Estimation of standby leakage power in CMOS circuits considering accurate modeling of transistor stacks", I*nternational Symposium on Low Power Electronics and Design*, California, pp. 239-244, 1998.

CHAPTER 3

STRAINED SILICON Vs SILICON: DEVICES

3.1 Introduction

As discussed in the previous chapter, new technologies have evolved which complement the advantages of scaling [1]. Moreover, the design techniques outlined in the previous chapter have been used to improve the yield of the circuits against the effects of process variation. These design techniques and the technology aids to improve the performance of the whole system were considered to be different entities. However, to extract better performance from a given technology node or the use of novel materials there is a need to optimise circuit design around specific details of the technology and the materials considering the device geometry, drive current, substrate leakage, and speed of the individual devices.

During the last decade the advancement of strained-Si growth and device performance [2] has attracted much attention for novel band-gap engineered hetero-structure devices with significantly improved performance compared to their conventional Si counterpart [3]. In the literature [3] strained-Si is predicted to be a future alternative to Si due to its enhanced performance as outlined in the previous chapter. One of the added advantages of incorporating strained-Si in the main

stream semiconductor industry is its compatibility with the existing conventional Si process technology [4, 5]. However, the device level performance improvement in strained-Si has not yet been exploited by the design community. This chapter describes the implementation of both strained-Si and conventional Si structures and compares the performance firstly as basic devices and thereafter in Chapter 4, at circuit level, when used to construct primitive logic functions.

The potential improvement of strained-Si devices over conventional technology is obtained through enhanced carrier mobility resulting from induced strain in the channel region of MOSFET [5]. There are a variety of mechanisms used by semiconductor industry to induce strain in the channel of transistors. The range of performance enhancement achieved by different straining conditions and under different operating conditions is described in the literature [5-8]. The results discussed in section 3.3.2 show that both n- and p-MOSFETs based on strained-Si provides a higher drive current and increased transconductance than Si devices.

In undertaking the analysis of device and circuit behaviour when realised using both standard Si and strained-Si technologies, extensive use was made of a suite of Technology Computer Aided Design (TCAD) tools. These programs assist semiconductor technologists to examine, effectively, trade-offs such as saturation current and subthreshold current in MOSFETs, effects of modifying device structures etc. to obtain a better understanding of device behaviour [9]. Since TCAD tools are very important in performing the analysis of the device behaviour described in this chapter a brief awareness of the toolset is given in the next section before the analysis and comparison of the device and circuit behaviour for the different technologies is discussed.

3.2 TCAD Simulator: An Introduction

For a large number of transistors per chip and owing to the invention of new structures and materials for devices, verification of circuits is becoming more complex and the use of analytical simulators for studying and designing the transistors are therefore necessary. The commercially available TCAD tools for

semiconductor device and process simulations are MEDICI [9] and TSUPREM4 [13] from Synopsys Corporation, ATHENA and ATLAS from SILVACO, TCAD-studio etc. For circuit model extractions AURORA from Synopsys and UTMOST from SILVACO are the most commonly used tools.

MEDICI is a device simulator, which allows user to create a two dimensional structure of a semiconductor device, including the definition of oxide and silicon regions, doping profiles to simulate the current – voltage characteristics of the device. TSUPREM4 can be used to replicate the IC manufacturing process. These simulators can be used to model any semiconductor device. AURORA is general purpose optimisation tool for fitting analytical models to data and extract parameters for circuit simulation. The mathematical models used in the suite of TCAD tools comprise a set of fundamental equations, for example, Poisson, the continuity and transport equations [2]. Poisson's Equation relates variations in electrostatic potential to local charge densities. The continuity and the transport equations describe the way that the electron and hole densities evolve as a result of transport, generation, and recombination processes. A number of physical models are incorporated into the simulator for accurate simulations, including the models for carrier recombination, photo generation, impact ionization, band-gap narrowing, band to band tunnelling, mobility and carrier life time [9]. The TCAD tool suite has the capability to analyse current density, electric field temperature etc. MEDICI features include transient and AC small-signal analysis, impact ionisation, gate current and ionisation integrals [9]. There are different transport models available in device simulators for example, drift-diffusion and hydrodynamic models. In this thesis for the device simulations using TCAD, hydrodynamic simulations were preferred over conventional drift-diffusion simulation due to the degradation of accuracy when smaller feature sizes for the devices are considered [2]. However, the accuracy also depends on the mesh density used to define the device. A very dense mesh was defined in places of interest in the device for example, the channel of the transistor. There is a variety of mobility models provided by the simulator which need to be selected and this selection was done during the process calibration. The calibration process is discussed in the sub-sections.

The interaction of TCAD tools of interest is shown in the Figure 3.1.

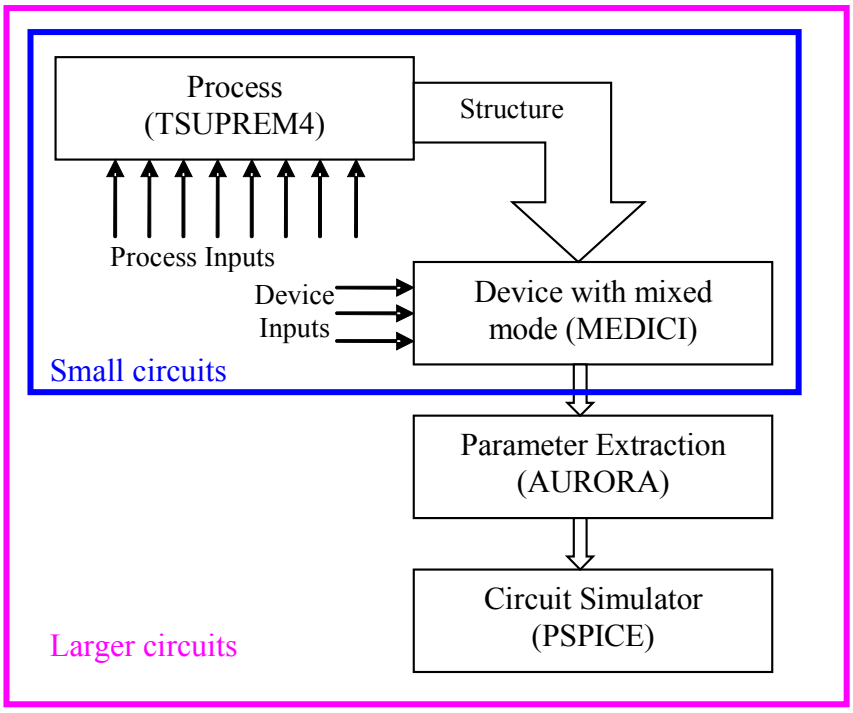


Fig. 3.1 TCAD tools [13].

As shown in Figure 3.1, if a larger circuit has to be analysed (For example, Jamb latch and MUTEX) then the mixed mode simulation could not be used due to the limitation on the number of nodes used to define the device. For larger circuits, the device characteristics is passed to the AURORA program which would generate the appropriate device models for a circuit level simulator such as PSPICE, the extracted models are shown in Appendix A. To demonstrate the advantages of strained-Si over conventional Si a CMOS inverter was designed in both technologies and its performance analysed. Since the inverter is a relatively simple circuit comprising an n- and a p-type transistor, a mixed mode simulator combining both the device and circuit level simulators, was used in the analyses. In this instance, the mixed mode approach was less time consuming and produced more accurate simulation results. The TCAD tools suite was used extensively for technology characterisation described in this chapter; for the analysis of circuit performance outlined in Chapter 4 and, finally for yield and variability analysis described in Chapter 5.

3.2.1 Method of simulator calibration

For accurate device characterisation the process and device simulators in the TCAD suite of tools must be calibrated. The calibration procedure for the process simulator comprises creating a virtual two dimensional device structure by selecting the appropriate process coefficients and process models, for example diffusion and ion implantation in the simulator. Thereafter the very tedious task is undertaken of 'matching' various profiles in the virtual device with those from the actual device; for example, doping profile extracted from a physical device using Secondary Mass Ion Spectroscopy (SIMS) [14]. The matching procedure, which effectively calibrates the simulator for a given process, is a lengthy task due to the large number of undefined coefficients which can be tuned in the empirical based models in the process simulator and the length of the subsequent process simulation runs; the procedure is also iterative. As an exact match will be rarely obtained there will also be slight discrepancies between the structure generated by the process simulator and the approximately extracted SIMS and Transmission Electron Microscope (TEM) data [14]. A two dimensional mesh structure of a conventional Si MOSFET constructed using MEDICI is shown in Figure 3.2.

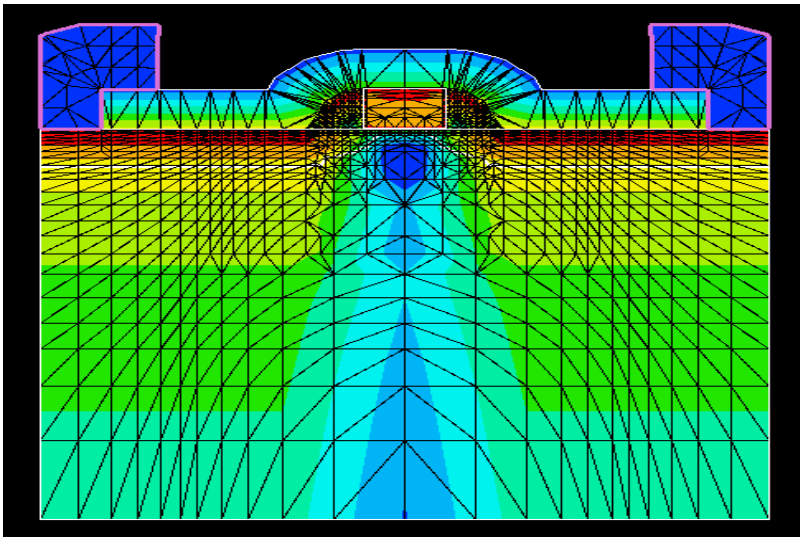


Fig. 3.2 Structure of a conventional MOSFET constructed using MEDICI.

The method of calibration adopted in this thesis is shown in Figure 3.3. The experimental data obtained using device electrical characterisation is compared with an example TCAD deck. The data included current voltage characteristics and dimension related parameters. The process of matching experimental data with example TCAD deck was continued until a near accurate match is obtained. In the matching process, different mobility models along with different doping profiles and dimensions were analysed. The mobility models available in TCAD are classified into two categories namely, Low field and Inversion layer mobility models.

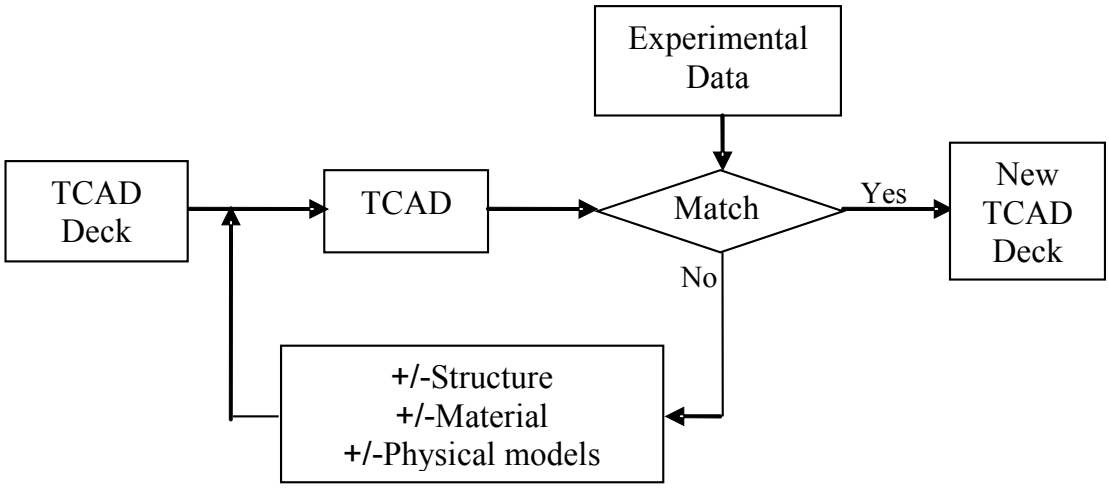


Fig. 3.3 Method of Calibration

Initially, mobility models namely, CONMOB, PRPMOB and FLDMOB which represents bulk, parallel and perpendicular fields respectively [9] were chosen due to the fact that operation of MOSFET's range from zero electric field to high electric field. To obtain a reasonable match of current voltage characteristics with that of the experimental data, doping profile, width of source-drain junctions etc were tweaked. For deep sub-micron devices, an enhanced mobility model SRFMOB2 which takes into account phonon scattering, surface roughness scattering and charged impurity scattering was selected. In the case of self-heating analysis, a temperature dependent mobility model TMPMOB was included in the TCAD deck along with hydrodynamic models [9]. The MEDICI input decks for 65, 90, 130 and 400nm devices which include doping profile, dimensions and physical models are shown in Appendix C.

3.3 Comparison of Strained-Si and Standard Si Devices

In comparing the characteristics of strained-Si and Si devices extensive use was made of the MEDICI simulator. The devices compared had gate lengths which spanned several technology nodes ranging from 400nm down to 45nm. The physical parameters for devices with technology nodes between 400nm and 130nm were similar to those fabricated at Southampton and Atmel. The specifications of deep sub-micron devices whose gate lengths were in the range 100nm to 45nm were taken from ITRS and MASTAR [11, 12]. The simulator was calibrated using ITRS and experimental data.

Figure 3.3 shows the structure of a strained-Si MOSFET considered for the analysis and comparison with conventional Si devices. In this work, we consider the strained-Si devices strained using bi-axial tensile strain. In the following discussion the percentage strain used is 0.99% which is obtained by including 25% Ge in the $Si/Si_{(1-x)}Ge_x$ (x=0 to 25%) virtual substrate layer (0.8μm thickness) in the device shown in Figure 3.3. However, the current/voltage analysis is performed for varying amount of strain in the channel which is discussed in the later sections of this chapter and continued in Chapter 4. The strained-Si channel considered here is 9nm thick. The virtual substrate is grown over graded $Si_{(1-x)}Ge_x$ which is placed on a substrate which has a uniform doping concentration of $1.5x10^{17}cm^{-3}$. The use of 25% Ge produces a stress equivalent to 1.8Gpa [15].

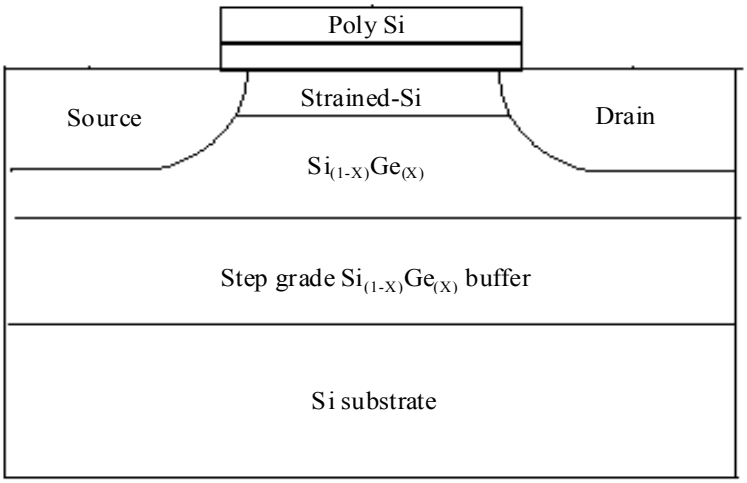


Fig. 3.4 Structure of a strained-Si MOSFET used for analysis.

Table 3.1 shows the amount of Ge percentage and the equivalent stress and strain (Hooke's Law) for the channel direction 100. Different versions of device structure similar to that shown in Figure 3.3 were developed depending on the variation of the dimensions and process parameters adhering to scaling rules.

Channel Direction (100)			
Ge %	**strained-Si (lattice constant)**	**Stress (in Gpa)**	**Strain (%)**
0 (Si)	5.430	0	0
5	5.441	0.357	0.2
10	5.452	0.715	0.4
15	5.462	1.072	0.6
20	5.473	1.430	0.8
25	5.484	1.787	1.0

Table. 3.1 Strain equivalent of the stress applied for different percentage of Ge in $Si/Si_{1-x}Ge_x$ material system [15].

Although there are different mechanisms available to induce strain in transistors, the enhancement of carrier mobility depends, only, on the amount of strain induced which will in turn improve the overall performance of the device [3, 7, 16, 17]. The enhancement of the mobility of the carriers in the channel leads to an increased flow of current in the transistor. The mechanism behind the enhancement of carrier mobility by inducing strain is discussed in detail in the sub section 3.3.1.

3.3.1 Mechanism behind strain

The transport properties of a semiconductor region are determined by the band structure which depends on the material, the composition of the material etc. Although there are different techniques for example, straining the channel of transistor, new device structures etc. to improve the transistor transport properties straining the channel is considered due to its processing cost advantages over other methods.

IBM has developed a method to strain the atomic structure of Si using $Si_{(1-x)}Ge_x$ alloys thereby obtaining a performance enhancement of about 70% compared to the conventional Si transistor [20]. A TEM image of the strained-Si transistor developed by IBM is shown in Figure 3.5.

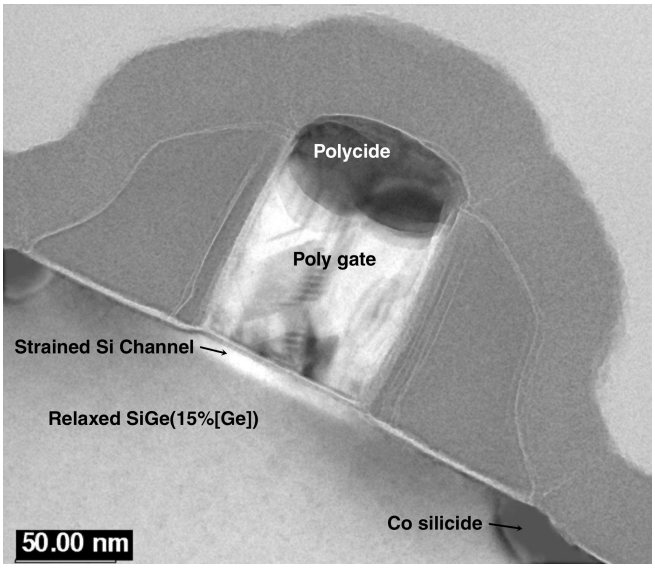


Fig. 3.5 TEM image of a strained-Si transistor developed by IBM [20].

Strain is the displacement of atoms in the crystal lattice from a normal position, the method by which this is achieved is as follows. When a Si layer is deposited on top of a substrate which whose atoms are spaced further apart as in the case of $Si_{(1-x)}Ge_x$ alloy, there is a natural tendency, as shown in Figure 3.6 for the silicon atoms to align themselves with the atoms in the substrate below, thus stretching or straining the Si layer. Conversely, when a thin layer of $Si_{(1-x)}Ge_x$ is grown on to a Si substrate the $Si_{(1-x)}Ge_x$ will be compressed.

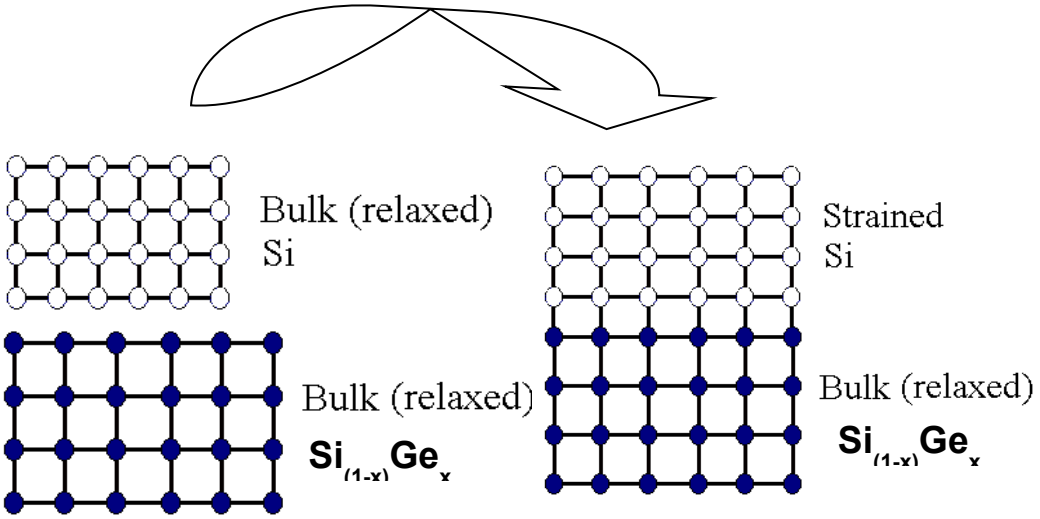


Fig. 3.6 Bulk Si grown on relaxed $Si_{(1-x)}Ge_x$ is strained because of different lattice constant [20].

The mechanism of strain and the transport properties of strained Si can be further explained using the band structure which is modified due to the high lattice mismatch between Si and $Si_{(1-x)}Ge_x$. Some of the material properties of interest of $Si_{(1-x)}Ge_x$ are compared with Si and Ge in Table 3.2 [5].

	Ge	$Si_{0.25}Ge_{0.75}$	$Si_{0.5}Ge_{0.5}$	$Si_{0.75}Ge_{0.25}$	Si
Minimum indirect energy gap (eV) (300K)	0.66	0.804	0.945	1.05	1.12
Lattice Constant (A)	5.6575	5.596	5.5373	5.4825	5.431
Thermal Conductivity (W/ cm- °c) (330K)	0.6	0.11	0.083	0.085	1.5

Table. 3.2 Material properties of $Si_{(1-x)}Ge_x$ compared with Si and Ge [5].

The band structure of Si is shown in Figure 3.7 where the conduction band minima is shifted (indirect band gap) in the <100> direction. The transport properties of Si are determined by this conduction band minimum and the valence bands where the split-off band is shifted by 44meV. The large hole effective mass and the close proximity of split-off band accounts for the poor transport properties of Si.

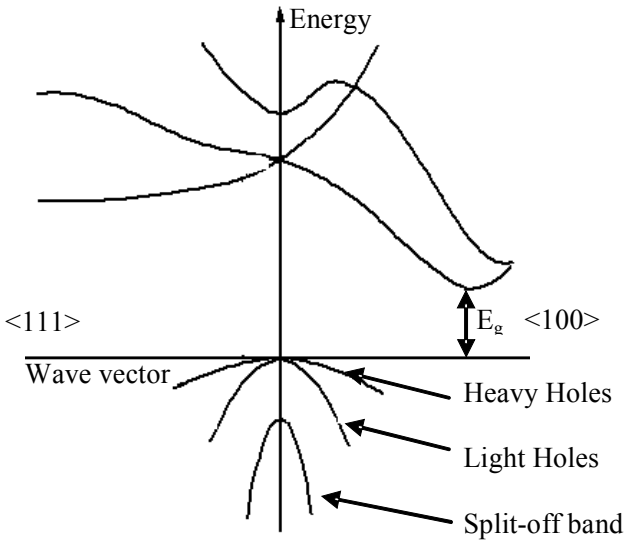


Fig. 3.7 Band Structure of Si.

The constant energy surfaces for the Si conduction band are shown in Figure 3.8. The six ellipsoids shown in Figure 3.8 are of the same size in unstrained Si with longitudinal (m_l) and transverse (m_t) electron effective masses of $0.98m_0$ and $0.19\ m_0$.

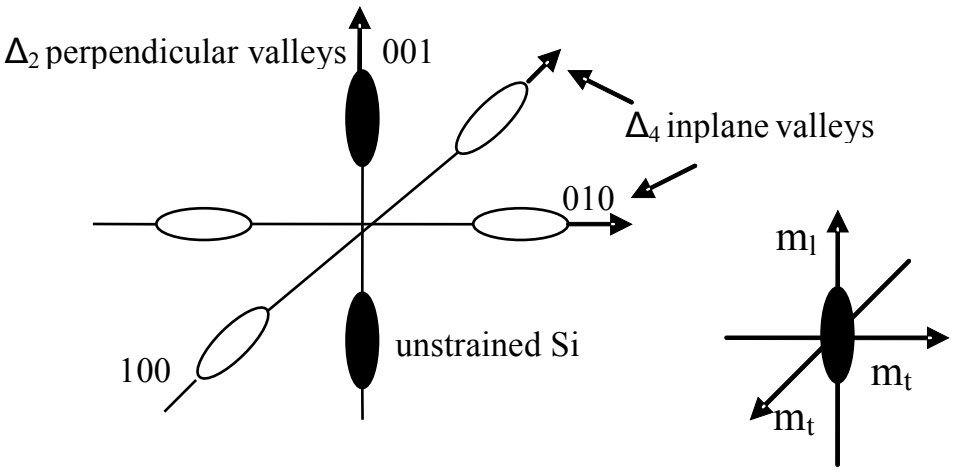


Fig. 3.8 Conduction band of Si.

When an epitaxial layer of Si is grown over $Si_{(1-x)}Ge_x$ the two fold degenerate Δ_2 perpendicular valley (001 band) will be shifted downwards in energy and Δ_4 inplane valleys will be shifted upwards in energy which is shown by the increase of the size of ellipsoid in the 001 band and decreased size in 010 and 100 bands as shown in Figure 3.9 (a). However, when a thin layer of $Si_{(1-x)}Ge_x$ is grown over thick Si substrate the four fold Δ_4 inplane valleys will be reduced in energy and the two fold Δ_2 perpendicular valley will be increased in energy as shown in Figure 3.9 (b).

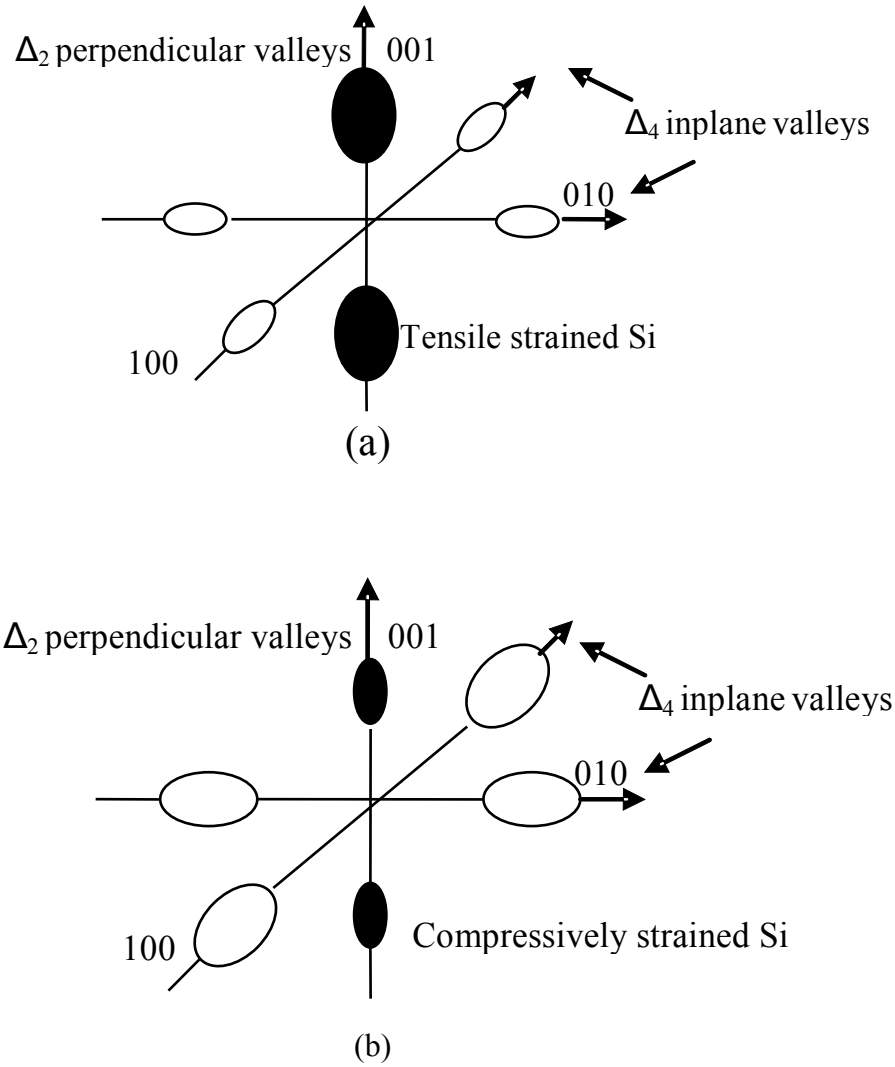


Fig. 3.9 Strained-Si conduction band energy states.

When the energy levels are reduced, the electrons occupy the lowest energy band which explains the improvement of transport properties of strained-Si. Valence band also contributes for the improved transport properties which is shown in Figure 3.10.

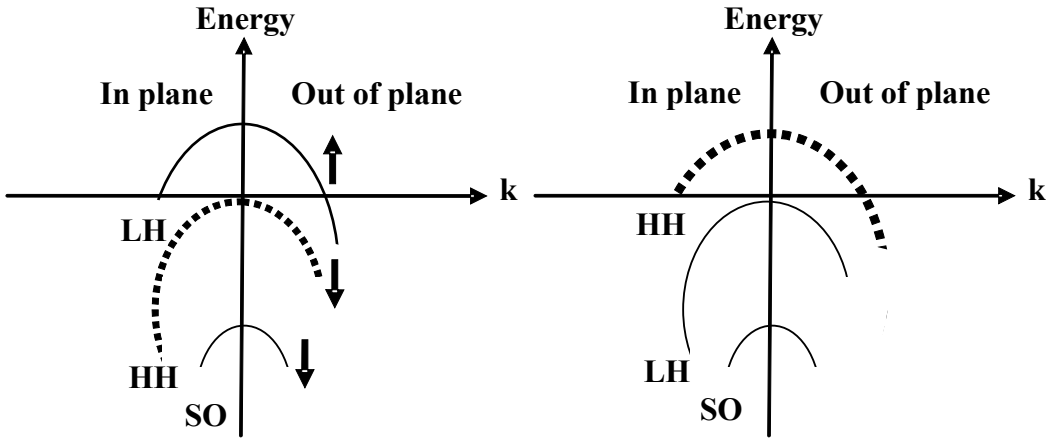


Fig. 3.10 Valence band Splitting in (a) tensile and (b) compressive strained Si

When strain is applied, it lifts the degeneracy between LH and HH bands which also changes the relative masses in the bands considerably which in turn improves the transport properties of strained-Si. In addition, the split-off bands is lowered in energy with respect to other two bands. These changes in the band structure for tensile and compressive strain can be classed as Type1 and Type 2 bands respectively.

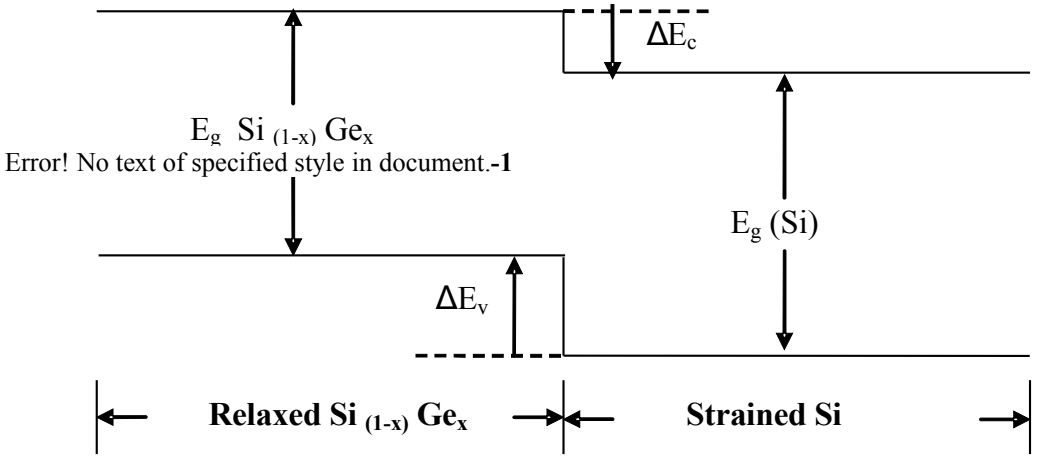


Fig. 3.11 Type 2 band alignment for tensile strain which improves electron mobility [5].

For Tensile stress, the conduction and valence bands will be aligned in a particular fashion called a Type 2 band alignment creating offsets in the valence

band and conduction band boundary between the two materials as shown in Figure 3.11. This band alignment effectively reduces the effective mass of the electron subsequently the electron mobility is increased as shown below [5].

$$\mu_n = \frac{q\tau}{m_n} \qquad \text{---------- (3.1)}$$

Where μ_n is the electron mobility

q is the electron charge

τ is the mean time between scattering events

m_n is the electron effective mass

Similarly, a Type 1 band alignment is for compressive strain where the band gap of the narrower material falls entirely within the forbidden energy gap of the wider band gap material creating offsets in the valence band and conduction band boundary between the two materials as shown in Figure 3.12. Here the offset in the conduction band is negligible compared to the offset in the valence band. This effectively reduces the effective mass of holes and hence the hole mobility is increased [6, 15, 16].

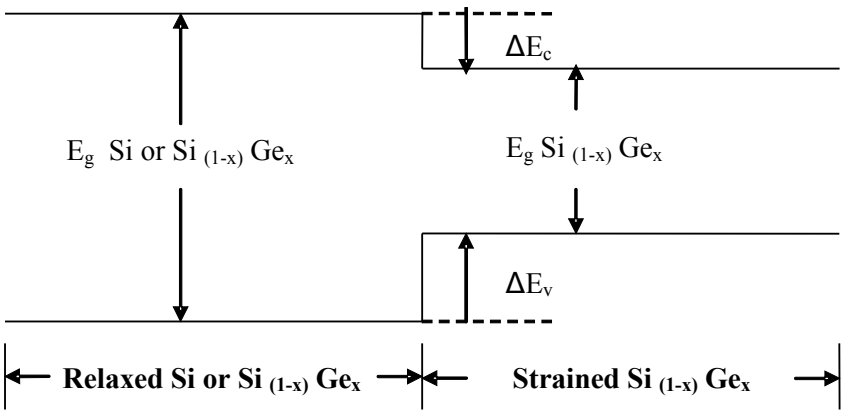


Fig. 3.12 Type 1 band alignment for compressive strain which improves hole mobility [5].

As discussed previously, tensile strain improves electron mobility (Si grown on $Si_{(1-x)}Ge_x$,) and compressive strain improves hole mobility ($Si_{(1-x)}Ge_x$ grown on Si). Hence tensile stress is required for n-MOS devices whereas compressive stress is required for p-MOS devices to enhance the performance [16]. However, from Figure 3.11 it should be observed that the offsets obtained in the conduction band side and the valence band side is large, making it ideal for both electron and hole mobility improvements. Also it should be noted that the offsets obtained depends on the concentration of Ge in the underlying $Si_{(1-x)}Ge_x$ which effectively determines the transport properties of strained-Si channel [5]. The enhancement of electron and hole mobilities achieved in industry is shown in Figure 3.13. There is a limit to the maximum thickness of Si which can be grown over $Si_{(1-x)}Ge_x$ to achieve the mobility enhancement and is called critical thickness which depends on the growth rates, growth temperature, and more on concentration of Ge. This method by which a Si layer is grown on $Si_{(1-x)}Ge_x$ is called wafer strain or global strain because the strain is induced uniformly across the wafer [4-6, 16].

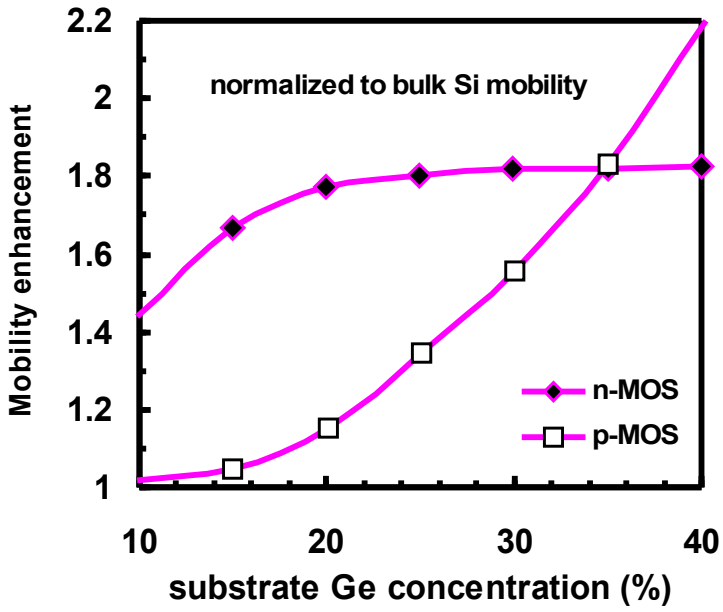


Fig. 3.13 The electron and hole mobility enhancements for different Ge concentration [21].

However, there are disadvantages with wafer level strain. One of the disadvantages is that the thermal conductivity of $Si_{(1-x)}Ge_x$ is less compared to the Si

layer above and hence the heat generated in the channel will not be easily dissipated through the substrate as in the case of conventional Si transistor and results in a phenomenon called self-heating [4, 5] which is discussed later in this chapter. To overcome this disadvantage, industry uses different methods to achieve strain in the channel of a transistor. Intel achieved the strain through process induced methods where, to stretch the silicon lattice, a film of silicon nitride is deposited over the whole transistor at high temperature. The advantage of using silicon nitride is that it contracts less than silicon as it cools; it locks the underlying silicon lattice in place with a wider spacing than it would normally adopt which is shown in Figure 3.14. Another method to introduce process induced strain is to make $Si_{(1-x)}Ge_x$ source and drain junctions on a Si substrate which is shown in the Figure 3.15.

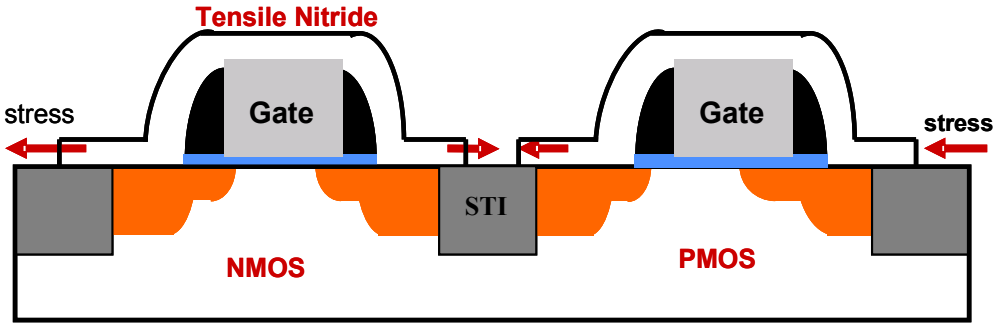


Fig. 3.14 Process induced strain achieved for n- and p-MOS by depositing a nitride layer on the top of whole transistor [22].

Generally speaking, no matter how the strain is induced the enhancement of carrier mobility depends on the amount of strain induced in the channel. Each method has its own advantages and disadvantages. In general one of the biggest advantages with strained-Si is that the stress methods can be varied in accordance to the circuit applications to achieve different amounts of mobility enhancement depending on where these devices are being used.

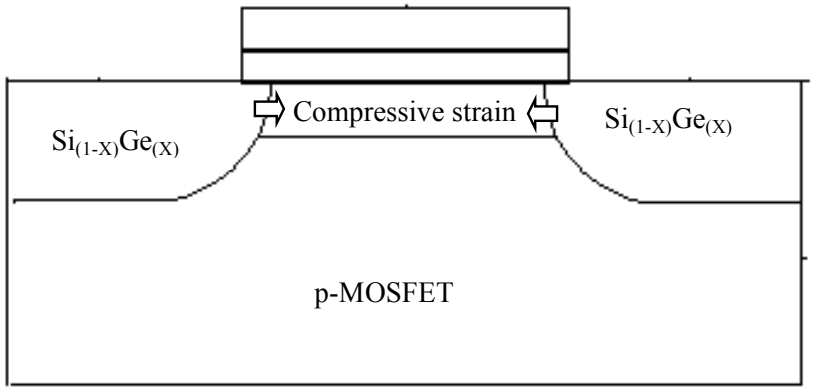


Fig. 3.15 Example of compression strain used by Intel [23].

3.3.2 Results and Discussion

The structure of a strained-Si MOSFET considered for the analysis and comparison with conventional Si devices has been discussed before and is shown in Figure 3.4. Although the particular devices considered in this work, to highlight the advantages of the technology, have a bi-axial tensile strain of 0.99% applied, the analysis, in general is performed for varying amounts of strain in the channel, as discussed in the later sections and continued in Chapter 4.

The parameters for the strained-Si/$Si_{1-x}Ge_x$ devices are derived from the reports [24], the key parameters used in the simulation are summarised in Table 3.3. Conventional Si MOSFET devices of similar dimensions were also constructed to compare their performance with that of the strained-Si devices. The low field and other mobility models including parallel and perpendicular fields with scattering have been incorporated into the simulation. The band gap narrowing model has also been considered. The low field electron and hole mobility values are based on reports [5, 16]. The electron affinity and other band parameters for $Si_{1-x}Ge_x$ are calculated using Vegard's Law [25].

Table. 3.3 Parameters for strained-Si/$Si_{1-0.25}Ge_{0.25}$ material system.

Device and material parameters	Value
Ge composition	25%
Strained Si channel thickness (nm)	9
Gate oxide thickness (nm)	3
Gate length (nm)	130 to 300
Junction depth (nm)	40
Channel doping (cm^{-3})	1.5×10^{17}
Substrate doping (cm^{-3})	1.5×10^{17}
Poly-Si doping (cm^{-3})	1.0×10^{19}
S/D doping (cm^{-3})	1.0×10^{20}
Density of states (cm^{-3})	$Nc = 1.0 \times 10^{19}$, $Nv = 6.0 \times 10^{18}$
Band offset modelling equation $Si_{1-x}Ge_x$ (x = 0 - 100) [24]	$\delta Ev = -0.238\ x + 0.03\ x^2$ -----(3.2) $\delta Ec = -0.35\ x - 0.35\ x^2 + 0.12\ x^3$ -----(3.3)
Low field mobility($cm^2/(V\text{-}s)$) [16]	Electrons=2340 holes=1000
High field mobility model [9]	PRPMOB, FLDMOB, SRFMOB2
Electron affinity $Si_{1-x}Ge_x$ (x = 0 - 100)	4.05 - 0.05 x (x = 0 – 100) -----(3.4)

The first step in the analysis is the calibration of the simulator using the drain current/drain voltage (I_{ds} - V_{ds}) experimental data available from Ghani *et al.* [26]. The gate dimensions and other physical parameters used for Si n- and p-MOSFETs are W/L = 1.0 μm/0.13 μm were also obtained from Ghani *et al*. [26]. Given that the calibration is performed to select the appropriate physical models to achieve a good agreement with the experimental characteristics, the same physical models are chosen for subsequent simulations. The simulations are compared with the available experimental data for conventional Si n-MOSFETs and are included in Figure 3.16 to verify the accuracy of the simulation. It is evident from Figure 3.16 that there is a

good agreement between sets of experimental and simulated characteristics for Si n-MOSFETs. P-MOSFETs have also been calibrated with available experimental data [26].

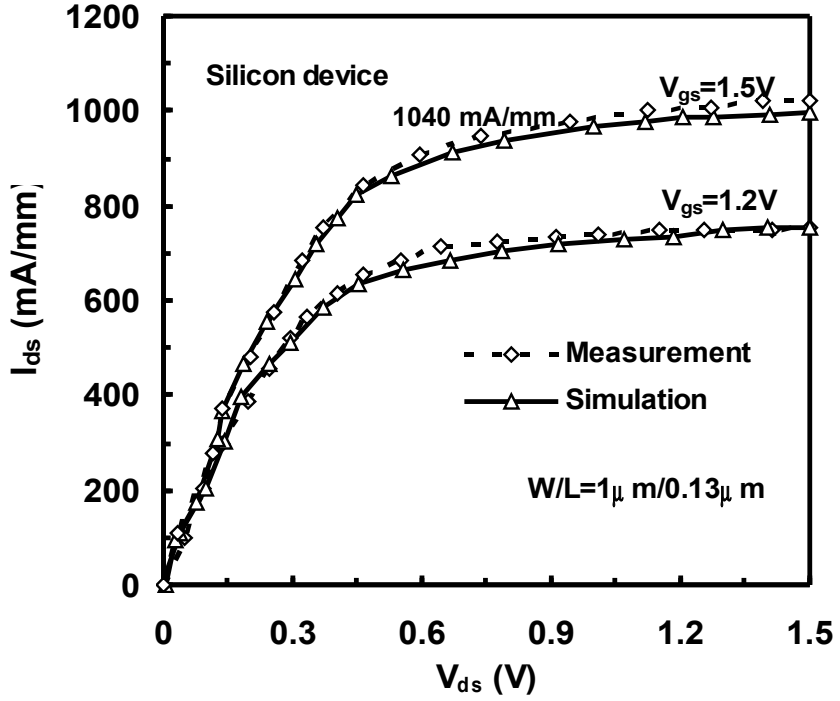


Fig. 3.16 Calibration of the $I_{ds} - V_{ds}$ characteristics for an n-MOSFET.

As a part of the calibration procedure, the experimental data using MOSFETs of different dimensions were also considered. Figure 3.17 shows the experimental subthreshold data for an n-MOSFET using the same physical models employed to calibrate the saturation region characteristics; the results shown are for a MOSFET in 400nm gate length. It is observed from the figure that the simulated results are in good agreement with the experimental data within an error of less than 2% which indicates the accuracy of the models and parameters used in the simulation.

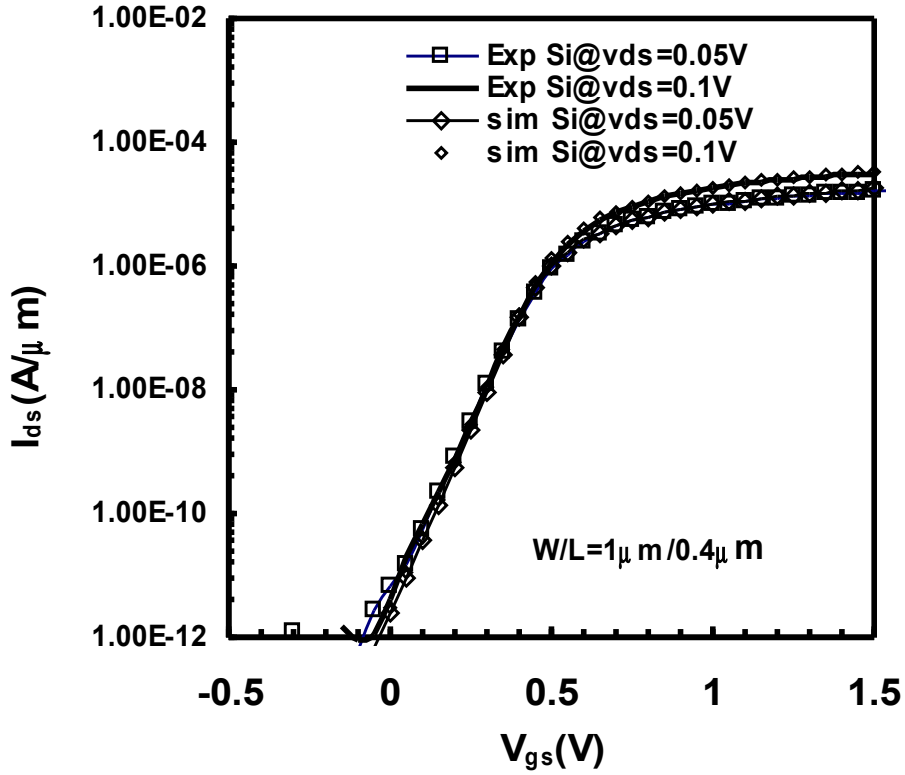


Fig. 3.17 Calibration of the subthreshold characteristics for an n-MOSFET.

Figure 3.18 and Figure 3.19 compare the drain current/drain voltage characteristics for strained-Si n- and p-MOSFETs respectively with conventional Si n- and p-MOSFETs having similar dimension for V_{gs} @ 1.2 V and 1.5 V. The devices have gate length of 0.13 μm, a Ge composition of 0.25 in the virtual substrate, and the gate oxide thickness of 3 nm. From Figure 3.16, it can be seen that the saturated drain current value for the Si n-MOSFETs is 1040mA/mm and 780mA/mm for V_{gs} @ 1.5V and 1.2V respectively. From Figure 3.14 it seen that for V_{gs} @ 1.5V the saturated drain current for the strained-Si n-MOSFET is 1250mA/mm indicating a 21% increase that the conventional Si counterpart. An increase of 28% in drain current is observed in case of strained-Si p-MOSFET compared to the Si p-MOSFET and is shown in Figure 3.19. The increase in drain current is attributed to the higher mobility of both electrons and holes in strained-Si compared to conventional Si as explained in section 3.3.1.

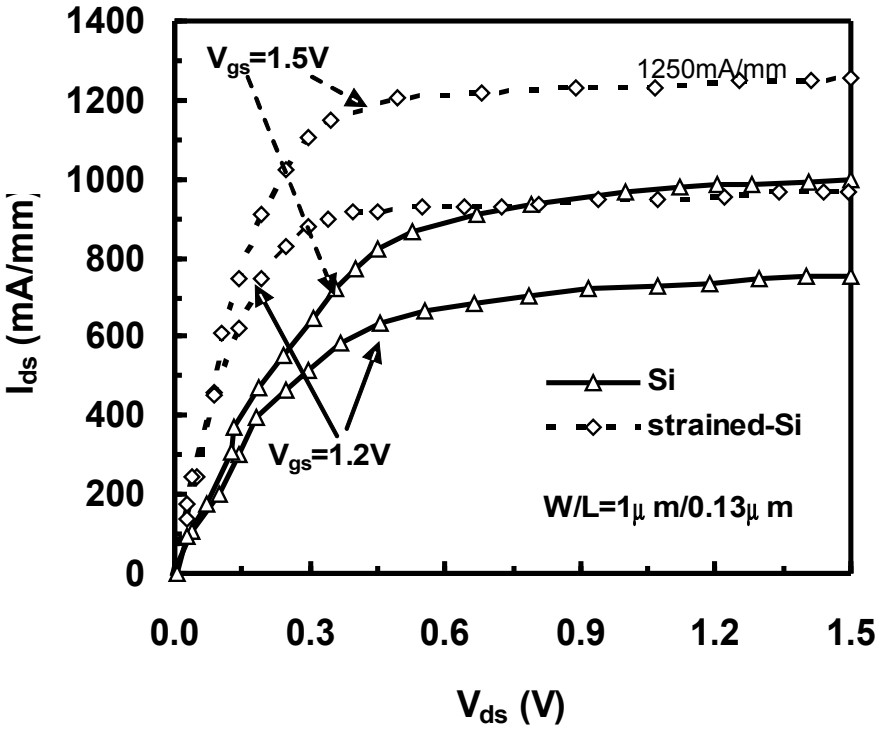


Fig. 3.18 Comparison of I_{ds} - V_{ds} characteristics for 0.13 μm gate length strained-Si and conventional Si n-MOSFET of similar dimension.

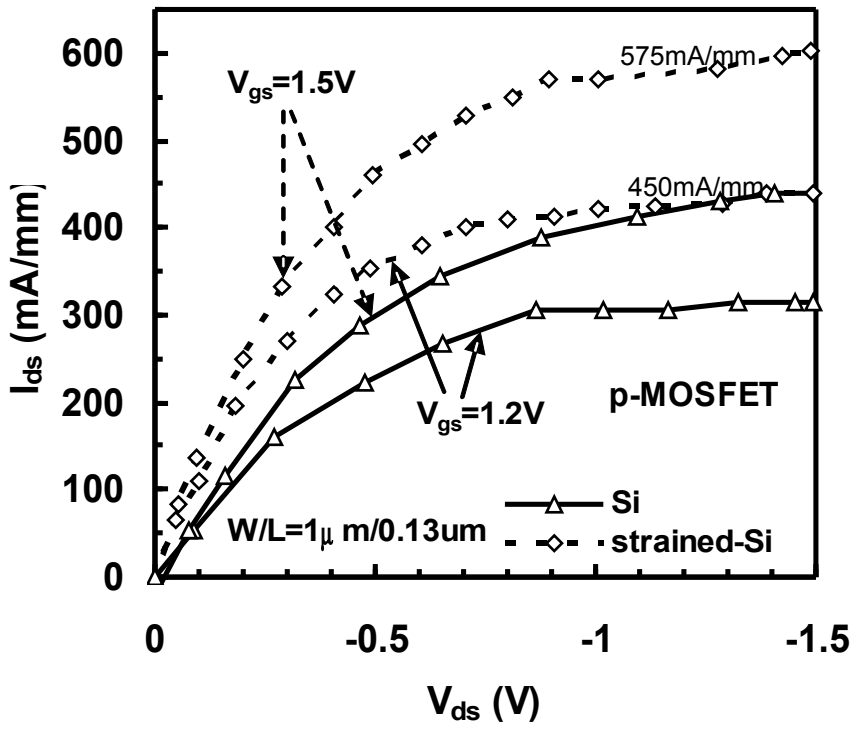
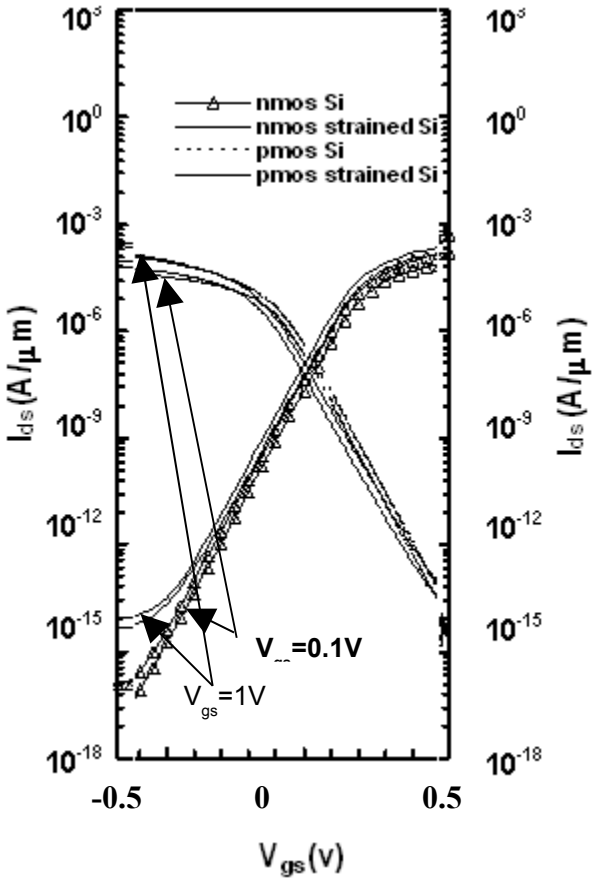


Fig. 3.19 Comparison of I_{ds} - V_{ds} characteristics for 0.13 μm gate length strained-Si and conventional Si p-MOSFET of similar dimension.

The drain current/gate voltage characteristics for strained-Si and conventional Si n- and p-MOSFET of gate length 0.13µm are plotted in Figure 3.20.

p-MOS



n-MOS

Fig. 3.20 Comparison of I_{ds} - V_{gs} characteristics for strained-Si and conventional Si n- and p-MOSFET devices with W/L = 1 µm/0.13µm.

It is interesting to note from Figure 3.20 that the leakage currents in strained Si n- and p-MOSFETs are almost three orders and one order of magnitude higher than in conventional Si n- and p-MOSFETs, respectively. The experimental data might exhibit a higher leakage current which is attributed to the presence of a greater number of defects and dislocations in the $Si_{1-x}Ge_x$ virtual substrate which are not included in the simulation. The subthreshold slopes for strained-Si n- and p-MOSFETs are found to be 91mV/dec compared to 80mV/dec in the conventional Si MOSFET. The threshold voltages for the strained-Si and conventional Si n-MOSFET devices are found to be 0.11 V and 0.30 V, respectively, and for p-

MOSFET devices these are -0.07 V and -0.29 V for strained-Si and conventional Si, respectively.

The transconductance of a MOSFET device is an important metric since it indicates the switching speed of the device. The plots in Figure 3.21 and Figure 3.22 compare the transconductance values for the strained-Si and conventional Si n- and p-MOSFETs respectively which is derived from Figure 3.20. It shows a peak transconductance enhancement of 100% in strained-Si device over conventional Si with similar dimensions. The higher transconductance obtained from strained-Si MOSFET devices indicate their suitability of for use in high performance circuits and this would be better suited than the conventional Si devices in this type of application.

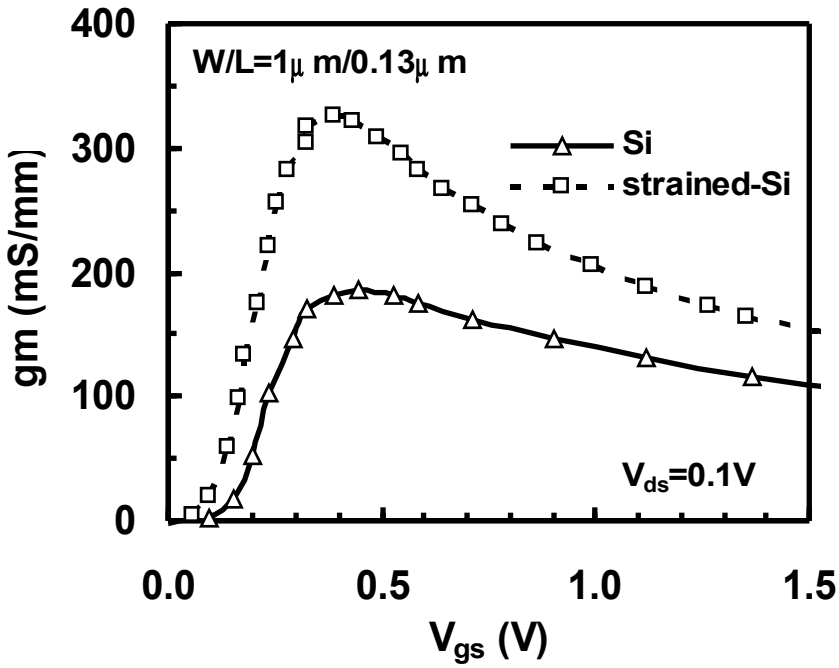


Fig. 3.21 Transconductance of strained-Si and conventional Si n-MOSFET device with a gate length of 0.13 μm as a function of gate voltage for V_{ds} @ 0.1V.

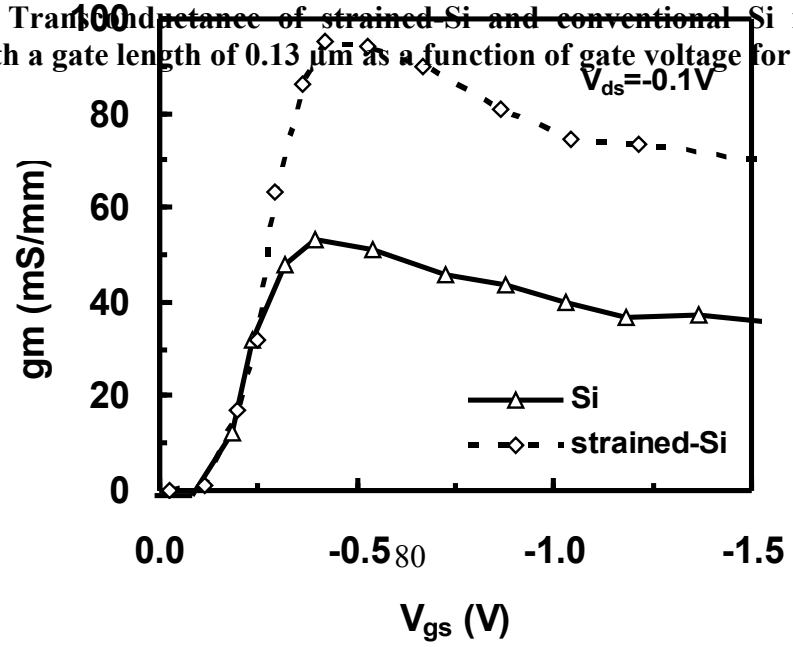


Fig. 3.22 Transconductance of strained-Si and conventional Si p-MOSFET device with a gate length of 0.13 μm as a function of gate voltage for V_{ds} @ -0.1V.

3.3.3 Self-heating

The results discussed in the previous sections demonstrate that strained-Si is a worthy contender to Si technology for future high–speed digital applications. However there are some disadvantages associated with the strained-Si technology which are now discussed. When the device dimensions shrink, lattice self-heating and hot-carrier transport effects become significant in their operation which leads to power dissipation in the channel [27, 28]. Besides, the presence of $Si_{1-x}Ge_x$, which has a lower thermal conductivity, in the device, increases the self-heating effects which eventually reduce the performance of the devices. The thermal conductivity of $Si_{1-x}Ge_x$ for varying Ge content is shown in Figure 3.23.

To analyse the self-heating effects, the hydrodynamic model [9] is used. The model includes lattice heating and the hot-carrier transport effects with simple local field-dependent relationships.

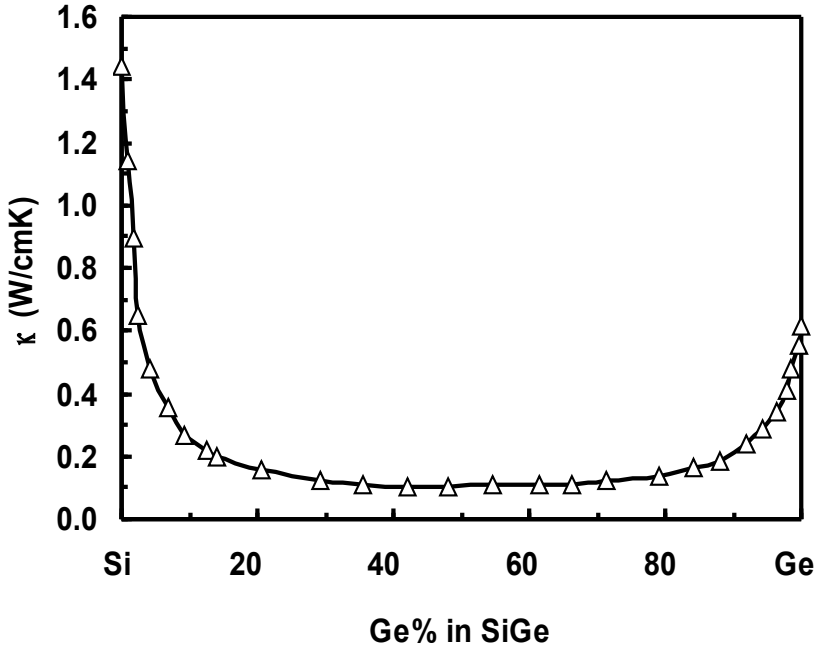


Fig. 3.23 Thermal conductivity of Si and $Si_{1-x}Ge_x$ for varying Ge content [5].

Figure 3.24 shows the DC current characteristics of an n-MOSFET device with self-heating obtained from simulation using MEDICI. The simulation results are compared with the experimental data obtained from a 400nm Atmel device. The device simulation is assumed to be at room temperature (ambient) for a good comparison with measured values. Analysis was performed to demonstrate how the self-heating affects the device operation

Power dissipated in the channel (self-heating effect) can be related to the difference between channel temperature and ambient temperature and the thermal resistance of the device. The Equation 3.5 shows this relation [29]:

$$\Delta T_c = P_t . R_{th} \quad \text{---------- (3.5)}$$

$$R_{th} = t_{SiGe} / K_{SiGe} A \quad \text{---------- (3.6)}$$

where R_{th} is the thermal resistance of the device; t_{SiGe} is the thickness of the underlying $Si_{1-x}Ge_x$; K_{SiGe} is the thermal conductivity of $Si_{1-x}Ge_x$ (which depends on the Ge content as shown in Figure 3.23) and A is the gate area (L_{eff} x W). For the

purpose of 2-D device simulation and calibration the underlying $Si_{1-x}Ge_x$ thickness is mapped to the thermal resistance. It should be noted that self-heating can only be observed when static DC measurements (when the frequency is set to zero) are used [14, 27, 29].

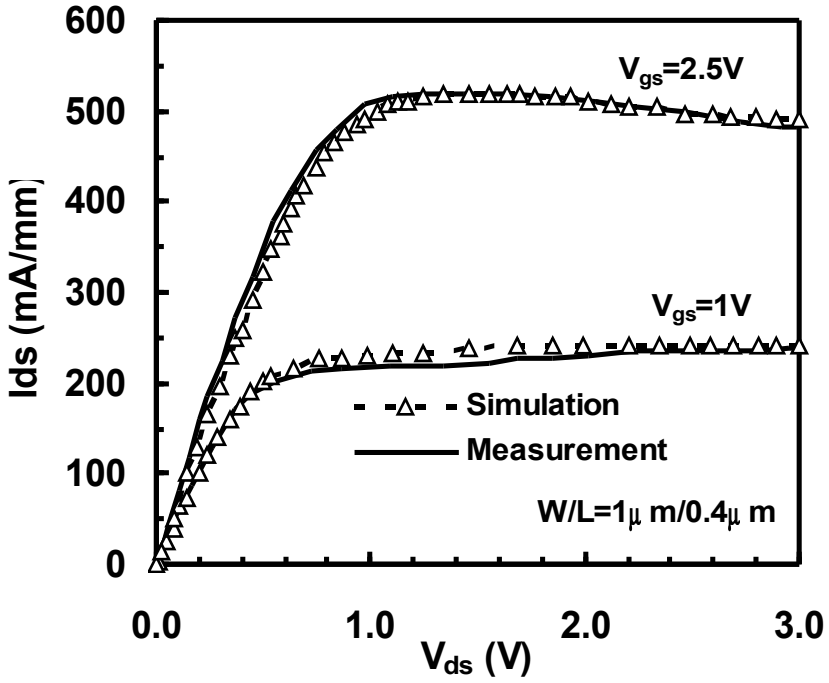


Fig. 3.24 Static Drain current Vs drain voltage characteristics.

It can be seen from Figure 3.24 that at a gate to source voltage (V_{gs}) of 1V, there is no reduction in current and is similar to the current/voltage characteristics discussed in section 3.3.2 and shown in Figure 3.18. However, at a gate to source voltage of 2.5V, the drain current starts to decrease in the strong inversion region (around 10% reduction at V_{ds}=3V). This phenomenon is due to the temperature accumulated in the channel and is called self-heating. Figure 3.25 shows the temperature rise in the channel against the increase of drain current. The increase of channel temperature accounts for the lower thermal conductivity (Figure 3.23) of underlying $Si_{1-x}Ge_x$ virtual substrate [27, 29]. The temperature accumulated in the channel will increase the kinetic energy of the carriers, which in turn will increase the scattering between the carriers, resulting in lower mobility, thus explaining the reason for the reduced current in the strong inversion region [27, 29].

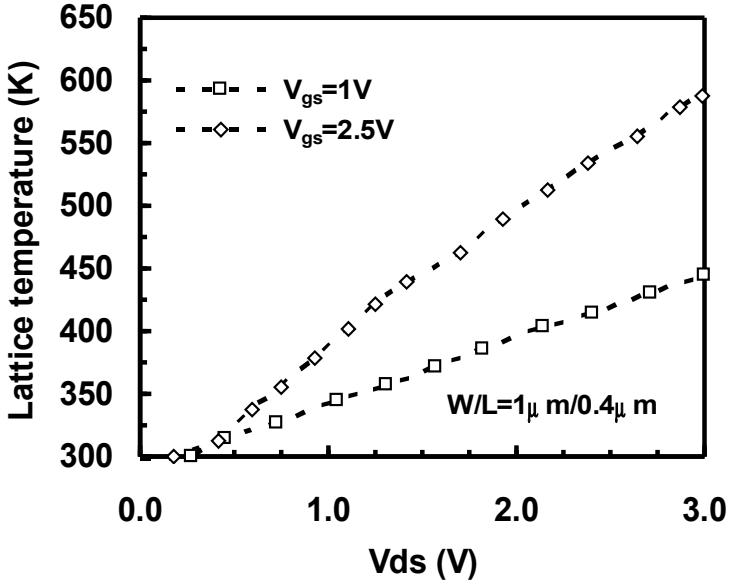


Fig. 3.25 The rise of temperature in lattice due to increase of drain voltage.

Figure 3.26 shows the transconductance of the device at different temperatures. The transconductance is an important parameter which determines the device speed [4, 5, 14, 29]. When the temperature increases, the transconductance is found to decrease as shown in the figure which in turn will decrease the device speed.

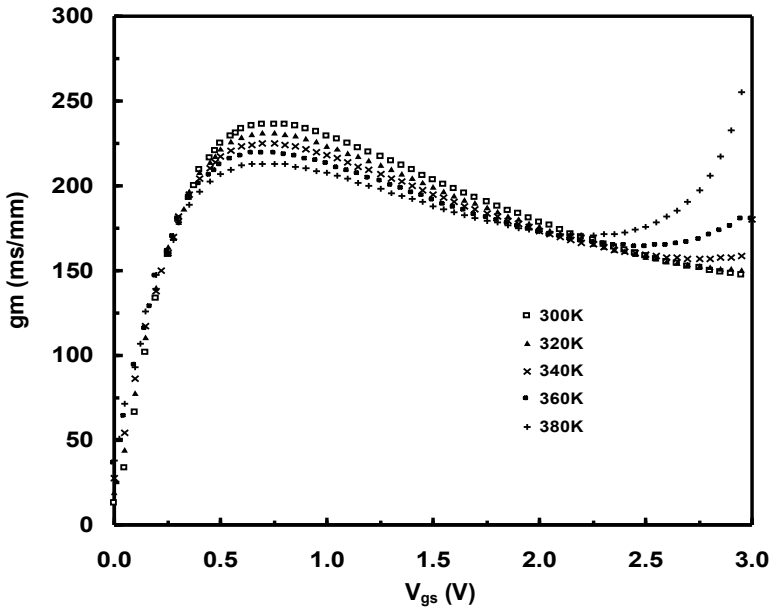


Fig. 3.26 Maximum transconductance against the temperature in the channel.

The variation of threshold voltage and leakage current against the ambient temperature in the channel is shown in Figure 3.27. The threshold voltage is found to decrease when the temperature increases and this is attributed to the fact that a greater number of carriers are generated in the channel which will reduce the voltage needed to form the channel. However, the leakage current is found to increase with respect to the temperature in the channel which is undesirable. These effects are important for devices in switching operation which are dependent on time. For continuously operated circuits, the increased leakage current due to reduction in threshold voltages is not significant [30].

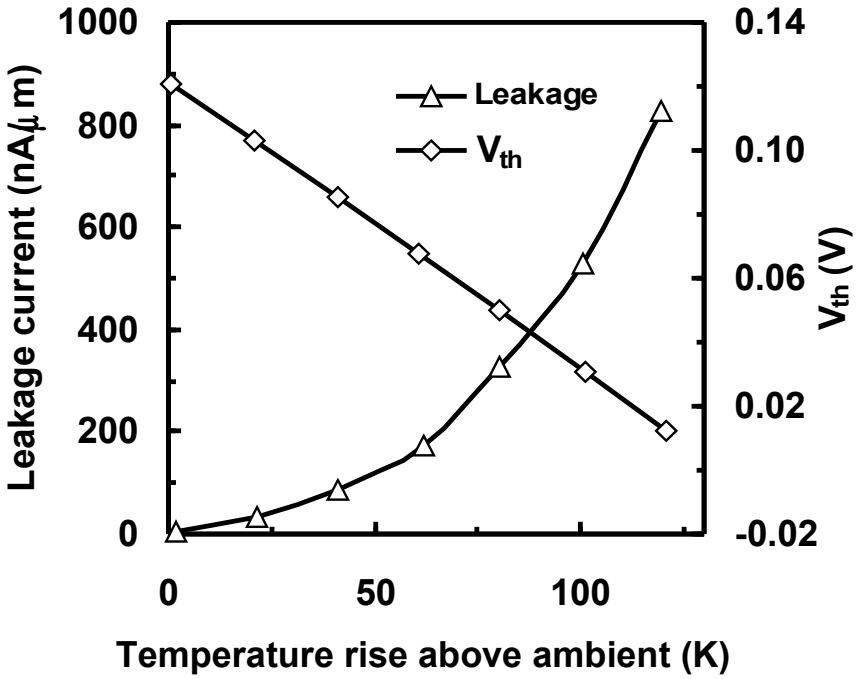


Fig. 3.27 Variation of threshold voltage and leakage current against the temperature rise in the channel.

Figure 3.28 shows the variation of mobility according to the temperature rise above ambient in the strained-Si channel. The values of mobility and temperature were taken by probing the channel near the drain end of the MOSFET using the utilities available in Taurus Work Bench (TWB). The mobility in the channel is found to decrease with the rise in temperature and the rise in temperature is due to the lower thermal conductivity of underlying $Si_{1-x}Ge_x$ virtual substrate as discussed in previous sections. The rise in temperature in the channel also depends on the

thickness of the underlying virtual substrate. Hence the increase of temperature can be controlled by changing the virtual substrate thickness which then reduces the effect of self-heating.

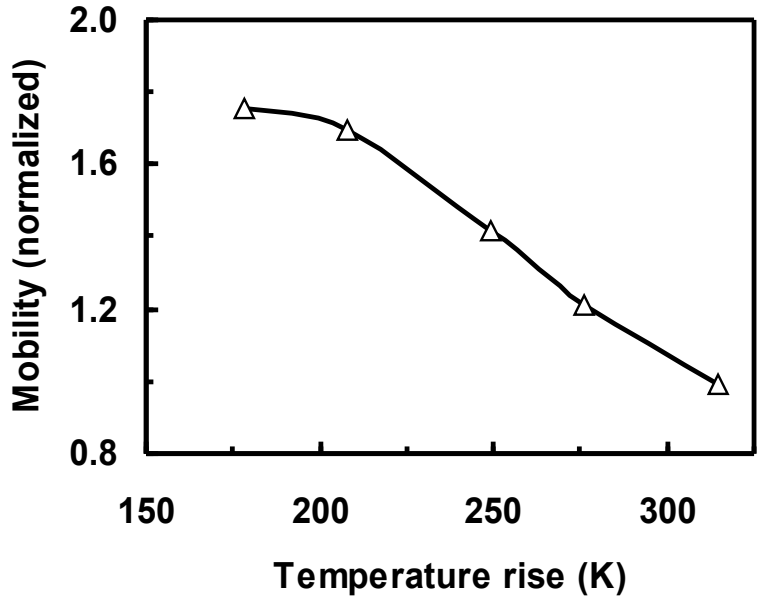


Fig. 3.28 Mobility degradation due to the temperature rise in the channel because of self-heating at higher channel voltages.

The percentage reduction in drain saturation current with respect to the current/voltage characteristics of strained-Si n-MOSFET shown in Figure 3.19 for normalized virtual substrate thickness is shown in Figure 3.29, where it can be observed that as the thickness of the virtual substrate increases the percentage reduction in the drain saturation current becomes larger. This is a result of the increase in the amount of self-heating due to increase in virtual substrate thickness decreasing the thermal conductivity. Thus it may be concluded that in order to minimise the reduction in drain saturation current due to self-heating, thin virtual substrates should be used. Furthermore from Figure 3.29, it can be observed that as the percentage of Ge in the virtual substrate is decreased the percentage reduction in drain saturation currents is also decreased; this can be attributed to the slightly higher conductivity [5] (as shown in Figure 3.24) in the virtual substrate. However, it should be noted that the mobility is reduced when Ge % in the virtual substrate is reduced which is discussed in section 3.3.1 and shown in Figure 3.13.

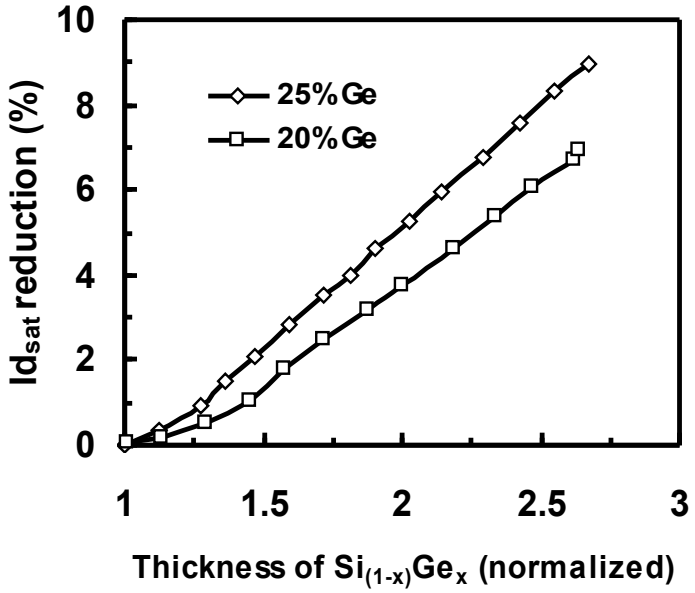


Fig. 3.29 The reduction of saturation drain current against the thickness of $Si_{(1-x)}Ge_x$ virtual substrate with 20% and 25% Ge content.

3.4 Summary

The TCAD tools used for the analysis of transistors have been reviewed and method of calibration has been discussed in this chapter. Strained-Si technology has been introduced and the physics behind it is outlined, thereafter the MOSFET characteristics based on bi-axial strained-Si is discussed and compared to that of conventional Si technology. The results indicate that strained-Si is a better technology for modern transistors with regard to saturation current and threshold voltage. The potential of strained-Si to combat scaling to some extent has been revealed through the results discussed. However, the inherent disadvantages associated with strained-Si technology based on b-axial strain that is, increase in leakage current and self-heating is discussed in detail. The physical parameters affected by self-heating are also discussed. It is seen from the results that self-heating becomes less important when thin virtual substrates are used in transistors.

It should be noted that, in order to exploit these advantages of strained-Si transistors in circuits, there is a need to optimise circuit design around specific

details of strained-Si technology considering the device geometry, drive current, substrate leakage, and considering the circuit application.

Having discussed the characteristics of individual strained-Si devices, the next stage is to analyse how these devices will perform in circuit level applications. Consequently in the next chapter, the most basic digital logic element namely, the CMOS inverter is constructed using strained-Si n-and p-MOS devices. Its behaviour, in terms of the important parameters, namely power consumption and speed is analysed. In addition, stacked inverters, ring oscillators, asynchronous circuit blocks like Jamb Latch synchroniser and Mutually Exclusion Element (MUTEX) are also developed using strained-Si technology and its performance are compared with conventional Si counterparts.

Overall:-

- Important aspects of TCAD simulators were reviewed.

- Different methods of achieving strain in the channel of transistors were discussed.

- Simulator is calibrated with experimental data within an error of less than 2% which indicates the accuracy of the models and parameters used in the simulation.

- The saturated drain current for the strained-Si n-MOSFET (0.99% strain in the channel) indicates a 21% increase that the conventional Si counterpart. An increase of 28% in drain current is observed in case of strained-Si p-MOSFET compared to the Si p-MOSFET.

- Subthreshold leakage current is seen to be higher by one order of magnitude for strained-Si devices compared to its Si counterpart.

- Self-heating which is a disadvantage with devices strained using wafer level method is discussed in detail.

- The transconductance variation due to self-heating is discussed.

- Variation of threshold voltage, subthreshold leakage current and mobility degradation due to self-heating was discussed.

- The reduction of saturation drain current against the thickness of $Si_{(1-x)}Ge_x$ virtual substrate with 20% and 25% Ge content was discussed to justify the use of thin virtual substrate for wafer level strained devices.

References

[1] M. Ieong, V. Narayanan, D. Singh, A. Topol, et.al., "Transistor scaling with novel materials", *Materials Today*, vol. 9, pp. 26-31, 2006.

[2] Y. J. Yang, W. S. Ho, C. F. Huang, S. T. Chang, et.al., "Electron mobility enhancement in strained-germanium n-channel metal-oxide-semiconductor field-effect transistors", *Applied Physics Letters*, vol. 91, pp. 102103 (1-2), 2007.

[3] Z. Cheng, J. Jung, M. L. Lee, A. J. Pitera, et.al., "Hole mobility enhancement in strained-Si/strained-SiGe heterostructure p-MOSFETs fabricated on SiGe-on-insulator (SGOI)", *Semiconductor Science and Technology.*, vol. 19, pp. L48-L51, 2004.

[4] S. Chattopadhyay, L. D. Driscoll, K. S. K. Kwa, S. H. Olsen, et.al., "Strained Si MOSFETs on relaxed SiGe platforms: performance and challenges", *Solid-State Electronics*, vol. 48, pp. 1407-1416, 2004.

[5] C. K. Maiti, N. B. Chakrabarti, and S. K. Ray, *Strained Silicon Heterostructures: Materials and Devices*. London: Institution of Electrical Engineers, 2001.

[6] R. L. Peterson, “Stretching silicon: a uniaxial and biaxial strain generation process and the resulting mobility enhancement in silicon-on-insulator MOSFETs”, Thesis (Ph. D.), Princeton University, 2006.

[7] L.J. Huang, J. O. Chu, S. A. Goma, C. P. D Emic, et.al., "Electron and Hole Mobility Enhancement in Strained SOI by Wafer Bonding", *IEEE Transactions on Electronic Devices*, vol. 49, pp. 1566-1571, 2002.

[8] V. Chan, K. Rim, M. Ieong, S. Yang,et.al., "Strain for CMOS Performance Improvement", *IEEE Custom Integrated Circuits Conference.*, vol. 2005, pp. 667-674, 2005.

[9] MEDICI user manual: *Two-Dimensional Device simulation program*, Sunnyvale: Synopsys Inc, 2000.

[10] Semiconductor Parameter Analyser User Guide: *Agilent Technologies 4155C/ 4156C*, California: Agilent Technologies Inc, 2004.

[11] International Technology Roadmap for Semiconductors, "Process Integration, Devices and Structures", 2007 Ed.

[12] Model for assessment of cmos technologies and roadmaps, *MASTAR,* ST Microelectronics*,* Advanced devices research group, 2000-2005.

[13] TSUPREM4 user manual: *Two-Dimensional Process simulation program*, Sunnyvale: Synopsys Inc., 2000.

[14] J.H. Liao, "Characterization of strained silicon", Thesis (Ph. D.), Arizona State University, 2006.

[15] Y. L. Tsang, S. Chattopadhyay, S. Uppal, E. A.-E.-C. Escobedo-Cousin, et.al., "Modeling of the Threshold Voltage in Strained-Si CMOS Architectures", *IEEE Transactions on Electron Devices*, vol. 54, pp. 3040-3048, 2007.

[16] S. E. Thompson, G. Sun, Y. S. Choi, and T. A. N. Nishida., "Uniaxial-process-induced strained-Si: extending the CMOS roadmap," *Electron Devices, IEEE Transactions on*, vol. 53, pp. 1010-1020, 2006.

[17] Z.-Y. Cheng, M. T. Currie, C. W. Leitz, G. Taraschi, et.al., "Electron Mobility Enhancement in Strained-Si n-MOSFETs Fabricated on SiGe-on-Insulator (SGOI) Substrates", *IEEE Electron Device Letters*, vol. 22, pp. 321-323, 2001.

[18] D. Long, *Energy bands in semiconductors.* New York: Interscience Publishers, 1968.

[19] *The secrets of strained silicon.* New scientist, pp. 27, Jan. 2004.

[20] P. M. Mooney, "Materials for Strained Silicon Devices", *IBM reserach report- Material Science*, pp. 1-11, 2002.

[21] (2002) IMEC website, [Online]. Available: http://www.imec.be/wwwinter/processing/asd/activities/strained.shtml.

[22] (2002) Intel website [Online]. Available: http://www.intel.com/technology/silicon/si12031.htm.

[23] C. Wee, S. Maikop, and C. Y. Yu, "Mobility-enhancement technologies", *Circuits and Devices Magazine, IEEE*, vol. 21, pp. 21-36, 2005.

[24] L. Yang, J. R. Watling, R. C. W. Wilkins, M. Borici et.al., "Si/SiGe heterostructure parameters for device simulations", *Semiconductor Science and Technology.*, vol. 19, pp. 1174-1182, 2004.

[25] K. T. Jacob, S. Raj, and L. Rannesh, "Vegard's law: a fundamental relation or an approximation?", *International Journal of Material Research*, vol. 98, pp. 776-779, 2007.

[26] T. Ghani, S. Ahmed, P. Aminzadeh, J. A. B. Bielefeld et.al., "100 nm gate length high performance/low power CMOS transistor structure", IEEE Electron Devices Meeting, 1999. IEDM Technical Digest. International, 1999.

[27] K. A. Jenkins and K. Rim, "Measurement of the Effect of Self-Heating in Strained-Silicon MOSFETs", *IEEE Electron Device Letters*, vol. 23, pp. 360-362, 2002.

[28] J. Olsson, "Self-heating effects in SOI bipolar transistors", *Microelectronic Engineering*, vol. 56, pp. 339, 2001.

[29] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Fabrication and analysis of deep submicron strained-Si n-MOSFET's", *IEEE Transactions on Electron Devices*, vol. 47, pp. 1406-1415, 2000.

[30] S. G. Narendra and A. P. Chandrakasan, *Leakage in Nanometer CMOS Technologies*. New York: Springer, 2006.

CHAPTER 4

FEASIBILITY OF STRAINED SILICON TECHNOLOGY FOR ANALOG AND LOW-POWER HIGH-SPEED DIGITAL CIRCUITS

4.1 Introduction

The power consumption of digital circuits is one of the main concerns of the present day design community, particularly with the market dominance of mobile and portable systems and the lack of critical innovation in battery technology. The supply voltage scaling has been emerged as the preferred technique for power reduction due to the quadratic relationship between the supply voltage and active power dissipation (alpha-power law) [1]. However, the associated penalty comes in the form of drastic increase in circuit delay and degradation of noise margin [2]. State of the art designs are struggling satisfy the projections of the International Technology Roadmap for Semiconductors (ITRS) [3] for operations in the sub-volt supply range, whilst maintaining high-speed. However, current circuit design research with Si-based CMOS is seriously challenged by these criteria to maintain high-speed operation [3, 4].

In the previous chapters, strained-Si has been introduced and its advantages in comparison to conventional Si technology were discussed. The results from the

previous chapter mainly concentrated on device performance. Even though strained-Si is now included in the ITRS roadmap the device level performance of strained-Si has not yet been fully exploited by the design community in high-speed low-power circuits. This requires research with closer integration of devices and circuits.

To assess the potential improvement in performance of digital logic functions as a result of using strained-Si devices in their implementation, the static and dynamic characteristics of the most basic logic element, namely the inverter were analysed. The inverter was constructed using standard Si CMOS and strained-Si CMOS technologies. The static characteristics considered were noise margin and static power consumption; the dynamic characteristics were gate delay, dynamic power consumption and slew rate. To demonstrate further the effects of improved strained-Si device performance at the circuit level, a ring oscillator was implemented using strained-Si CMOS technology. The results of this analysis are discussed in detail in the next sections.

This chapter also explores the possibility of performance enhancement in terms of power along with speed using strained-Si devices for digital circuit design. Although in this chapter the studies are based on the general behaviour of digital CMOS circuits using strained-Si as the base material, emphasis is given to noise and power performance analyses. This stems from the fact that if a trade-off between speed and supply voltage is made to reduce active power, the most likely parameter that is affected is the noise margin [5]. The entire results reported in this chapter are done by studying the static and dynamic characteristics of CMOS inverter circuits subjected to different operating conditions and different band-engineered devices.

The chapter also addresses the impact of voltage scaling and band gap engineering on leakage current, speed and noise margin at different ambient temperatures, since they are crucial in any digital design space. The rest of the chapter is structured as follows:

- The performance of strained-Si inverters is studied under different straining conditions in terms of noise and power with reduction of supply voltage.

- Device stacking method is used to reduce further power consumption with the advantages of using strained-Si technology.
- The analog performance of strained-Si technology is evaluated by building CMOS inverter amplifiers.

The results show that even under the reduced supply voltage condition and the stacking of transistors, strained-Si based inverter shows improved performance in terms of speed and power consumption compared to Si based inverters.

4.2 Performance Analysis

Straining of the channel material in CMOS can be attained either at the wafer level by epitaxially growing a thin film of Si over $Si_{1-x}Ge_x$ alloy [6] for example, at the process level by depositing SiN_3 on top of the device structures [7] discussed in Chapter 3. While compressive strain is required to enhance the mobility for holes, tensile strain can enhance mobility for both electrons and holes [8]. For the analysis, the devices considered have high performance achieved by bi-axial tensile strain. No matter how the strain is induced (whether bi-axially or uni-axially), the enhancement of carrier mobility depends on the amount of strain induced which will in turn improve the overall performance of the device [7, 8]. The dimensions of the devices used for the analysis ranges from 130nm down to 65nm.

Inverters designed using strained-Si and Si CMOS were analysed for the performance improvement of strained-Si technology compared to the conventional Si technology. The analysis was performed to evaluate both the propagation delay (transient characteristics) and DC transfer characteristics for the inverters. Here the sizes of transistors are considered the same. However, since the mobility enhancement of electrons and holes is asymmetric for the same amount of strain [7], the symmetric DC characteristics can only be attained by using different n- and p-transistor geometry ratios.

4.2.1 Static Characteristics

In analysing the static characteristics of strained and standard CMOS circuits, the main attributes of interest are noise margins.

4.2.1.1 Noise Margin

When devices are used in building digital circuits, unwanted voltages (noise) can be internally generated at the nodes in a circuit and depends largely on the properties of transistors used in the circuit [5]. One of the critical requirements of modern digital circuits is its immunity to noise.

Noise immunity or noise margin is a metric to measure the robustness of a digital circuit against noise. The mapping of a voltage range (input/output) for a particular logic state gives the measure of noise margin of a digital circuit [9]. A range of output voltages is mapped to 'low' and 'high' states as shown in Figure 4.1.

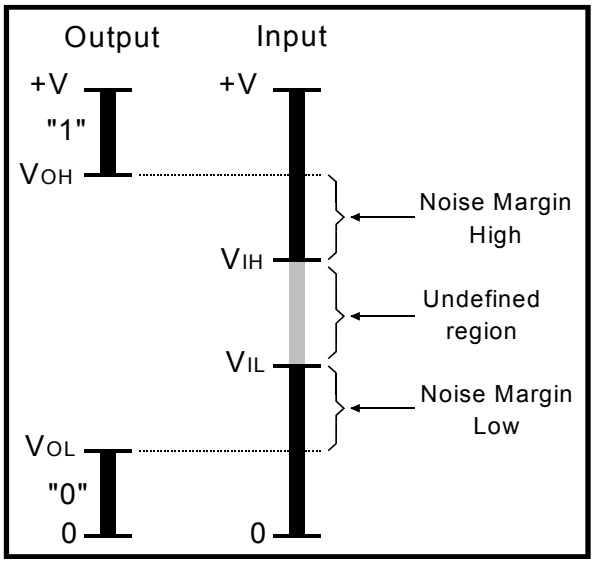


Fig. 4.1 Mapping of the voltage range for a particular logic state.

The maximum logic '0' input (V_{IL}), maximum logic '0' output (V_{OL}), minimum logic '1' input (V_{IH}) and minimum logic '1' output (V_{OH}) voltages are calculated from the slope of Voltage Transfer Curves (VTC). The voltage range

which cannot be mapped to a particular logic state, i.e., logic '0' or logic '1', is termed as the undefined region or metastability window and is calculated to be: undefined region (input range) = V_{IH}-V_{IL} [5].

The Noise Margin (NM) is the difference between what the driver IC outputs see as a valid logic voltage and what the receiver IC expects to see as a valid logic voltage. There are two different types of noise margin, one for a logic high value 1 and one for a logic low value 0. For a valid logic high, the worst case noise margin for the circuit is the minimum high level voltage, which may be output from the driver minus the minimum high level voltage that may be seen at the receiver IC [5]. Thus the Noise Margin Output high equals V_{OH} [driving device] minus V_{IH} [receiving device]. For a valid logic low, the worst case noise margin for the circuit is the maximum low level voltage which may be output from the driver minus the maximum low level voltage which may be seen at the receiver IC; that is Low Noise Margin Output equals V_{IL} [receiving device] minus V_{OL} [driving device].

NM_L and NM_H can be calculated from the DC transfer characteristics (from the unity gain point; where gain A_v is -1) as shown in Equation 4.1 to Equation 4.4 [29]:

$$V_{IL} = V_M - (Vdd - V_M)/ A_v \quad \text{---------- (4.1)}$$

$$V_{IH} = V_M - V_M / A_v \quad \text{---------- (4.2)}$$

$$NM_L = V_{IL} - V_{OL} = V_M + (Vdd - V_M) / A_v \quad \text{----------(4.3)}$$

$$NM_H = V_{OH} - V_{IH} = Vdd - V_M + V_M / A_v \quad \text{----------(4.4)}$$

where V_{IL} = input voltage where slope of transfer characteristic is -1.

V_{IH} = larger input voltage where slope of transfer characteristic is -1.

V_{OH} = output voltage at input voltage of $V_{IL.}$

V_{OL} = output voltage at input voltage of $V_{IH.}$

V_M = voltage where output voltage equals input voltage.

The voltage transfer curves of inverters are used to obtain parameters such as noise tolerance, gain and operating logic levels [9]. The VTC of strained-Si and conventional Si CMOS inverters for a drawn gate length (L_g) of 130 nm are compared in Figure 4.2. The circuit supply voltage is 1.8V. The width of p-MOSFETs for both strained-Si and conventional Si used in designing the inverters is taken to be nearly twice the width of an n-MOSFET to account for higher electron mobility in comparison to the mobility of holes [5, 9] to provide a symmetric voltage transfer characteristic [5, 9]. From Figure 4.2 it can be seen that the VTC for the conventional Si inverter is comparatively more symmetrical than the VTC of corresponding strained-Si inverter. This is due to the difference in the increase of electrons and hole mobilities in strained-Si technology for the same amount of strain applied.

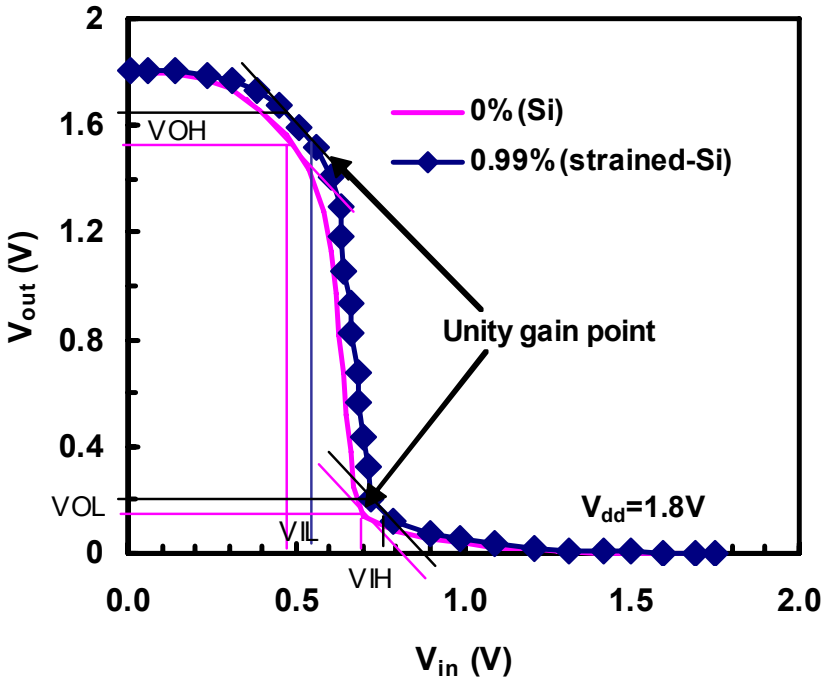


Fig. 4.2 Voltage transfer curve (VTC): Output voltage (V_{out}) as a function of input voltage (V_{in}).

It should be noted from Figure 4.2 that the output high to low and low to high transition unity gain points (where the transfer characteristic is '-1') [5] of strained-Si with 0.99% strain is different from that of Si inverter (0% strain). From the unity gain points the V_{IL}, V_{OL}, V_{IH} and V_{OH} voltages are noted by extrapolating

the points to the input and output voltage axis. Subsequuently, the noise margins of strained-Si and convention Si inverters are calculated as discussed previously.

Figure 4.3 shows the variation of low noise margin (NM_L) and high noise margin (NM_H) of a strained-Si inverter with 0.99% strain in the channel in comparison to Si inverter against supply voltage reduction. NM_L and NM_H are seen to degrade with supply voltage reduction which is expected, since the threshold voltage and 'input and output transition points' are scaled with the supply voltage. It should be noted that the NM_L for strained-Si inverter is higher due its inherent property of lower threshold voltage and higher drain current for strained-Si devices compared to Si devices [10]. However, the NM_H of strained-Si is roughly same as that of the Si indicating that the overall noise margin for the strained-Si CMOS inverters are better in comparison to similar conventional Si counterpart. However, this also depends on the width ratio of transistors.

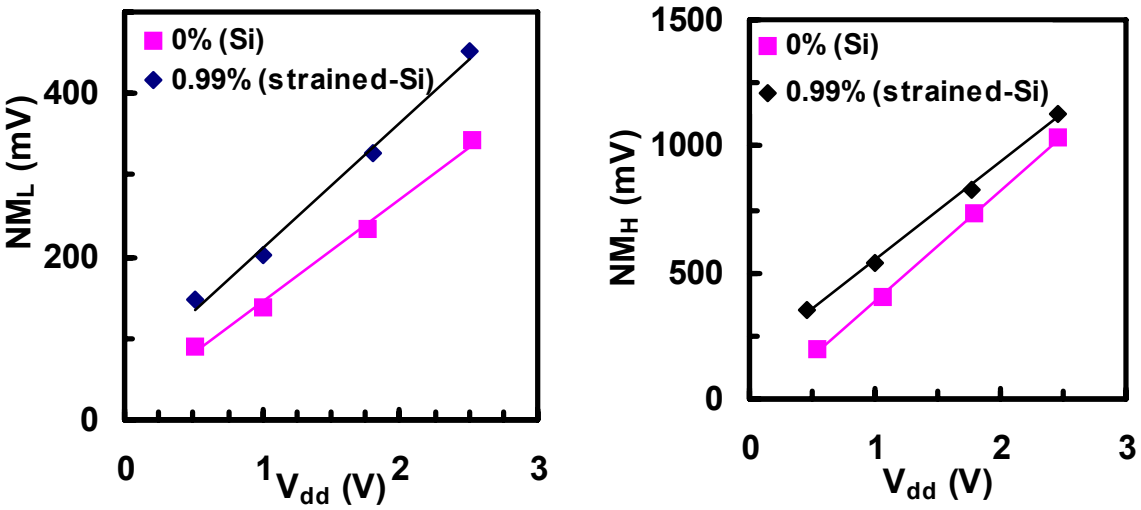


Fig. 4.3 Variation of NM_L and NM_H with supply voltage scaling for different strained-Si and Si CMOS inverters.

Figure 4.4 shows the variation of NM_L and NM_H for strained-Si inverters with different amounts of strain in the channel of the transistors. It should be noted that NM_L is increasing whereas NM_H is decreasing with increase in the amount of strain in the channel. This is due to the fact that, with the increase of amount of

strain, the mobility of the carriers in the channel are increased which will increase the process transconductance (K) which will in turn increase the device transconductance (β) [11-13]. With the increase in the amount of strain, the mobility enhancement for electrons and holes are different [7, 14] (it is higher for holes at the higher amount of strain) and hence the transconductance ratio (βn/βp) decreases for same n- and p-transistor geometry ratios and the voltage transfer curve shown in Figure 4.2 shifts towards right which explains why NM_L is increasing with increasing amount of strain while NM_H decreases. Hence bi-axial strain can be optimized for symmetrical VTC depending on the mobility enhancement.

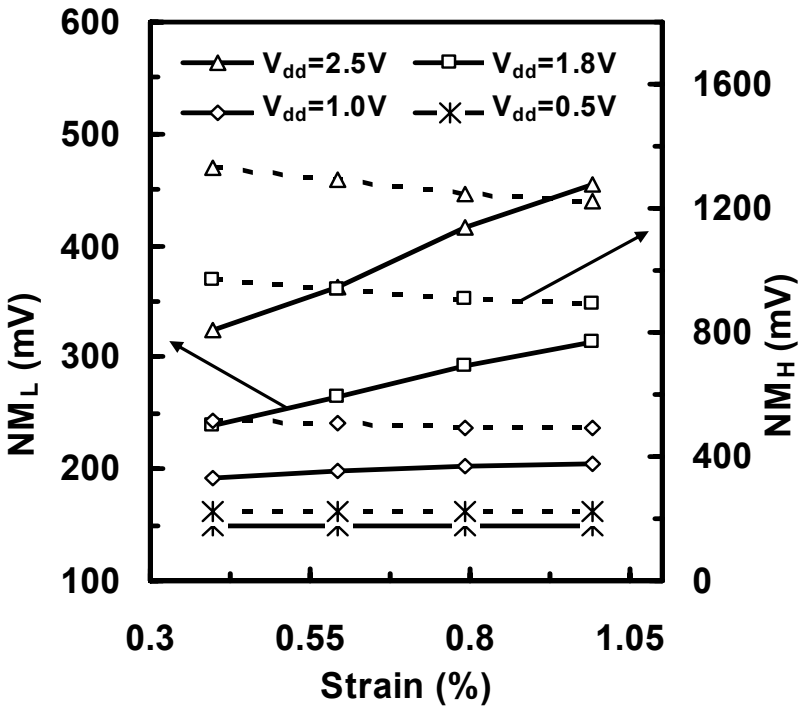


Fig. 4.4 The variation of NM_L (*solid lines*) and NM_H (*dotted lines*) with the amount of induced strain in the device channel at different supply voltages.

Another interesting fact is that as the supply voltage is lowered below 1V (which is the typical operating condition for the deep sub-micron circuits) the NM_L and NM_H show very little dependence on strain although at higher supply voltages they show significant dependence on the strain. From this result, it seems that when sub-1V supply is applied, the amount of strain present in the channel matters very little as far as the noise immunity is concerned. Thus, for sub-volt operation the optimization of strain should be more focussed on optimizing speed and power only.

The operation of inverter circuit was also analysed for diffferent ambient temperatures and the plots of noise margins are shown in Figure 4.5. The decrease of noise margins shown is attributed to the reduction of current gain (I_{on}/I_{off} ratio) for the degraded threshold voltages and subthreshold slopes of the transistors at elevated temperatures

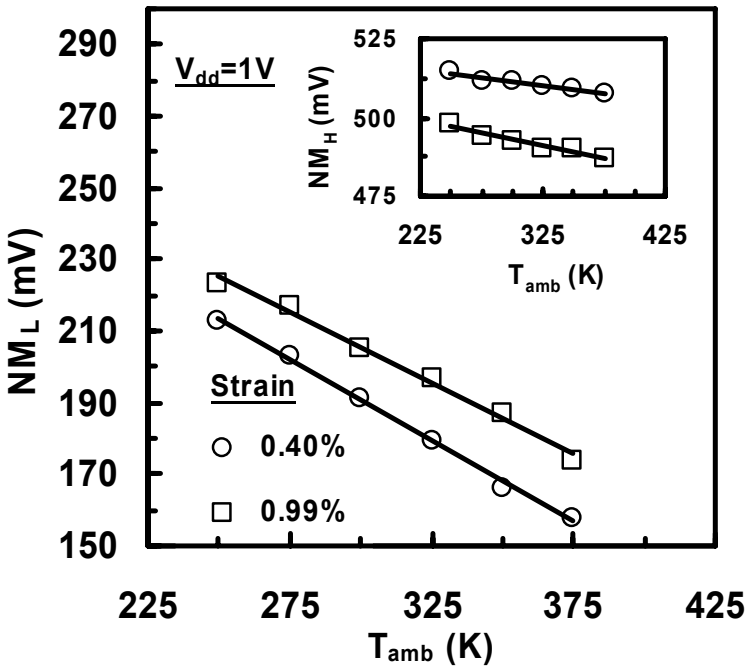


Fig. 4.5 The variation of NM_L and NM_H with the ambient temperature @ V_{dd} = 1V for different amount of strain along the device channel.

In summary from the above results the degradation of noise margins exhibited by strained-Si circuits is not worse than Si circuits when the supply voltage is reduced to achieve low power performance. On the contrary, strained-Si based circuits show significantly better noise margin compared to Si based circuits under the similar operating conditions.

Figure 4.6 shows the comparison of the undefined regions in the input voltage ranges of strained-Si and conventional Si CMOS inverters with different device gate lengths. The undefined region in the input voltage range in strained-Si devices is found to be 27 % less than the conventional Si at a gate length of 130 nm. This is attributed to the higher drive current and lower threshold voltage of strained-Si devices compared to bulk Si devices [6, 8, 15-18]. The undefined region is almost constant with gate length variation as opposed to that of conventional Si CMOS and

is useful when these devices are used to build latches and flipflops where metastability is a crucial factor. If the voltage given to the next stage of a sequential circuit falls into this metastability window then there will be intermittent failures in the operation of this circuit. The strained-Si CMOS devices therefore indicate that they are better candidate in sequential circuit applications as the probability of strained-Si CMOS based sequential circuit violating set-up time and hold-time requirements (i.e. metastability) is less.

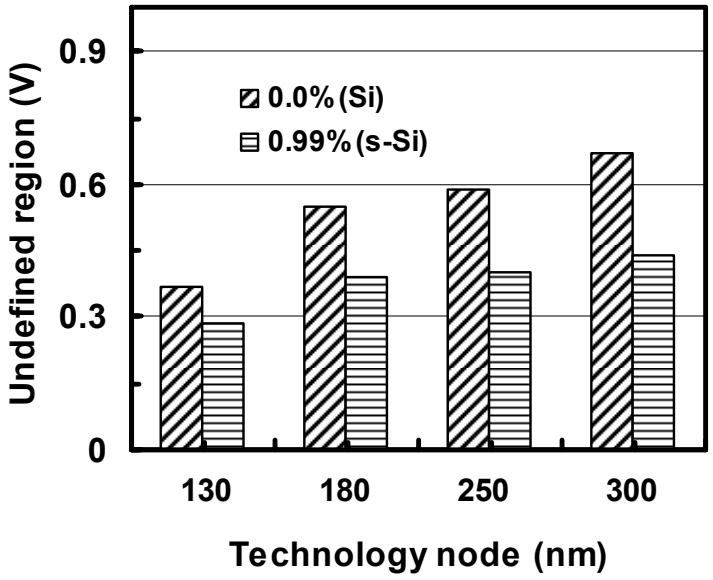


Fig. 4.6 Undefined region of strained-Si inverters against conventional Si CMOS inverters for different technology nodes.

Consequently a smaller undefined region for strained-Si inverter is deemed to be profitable from the noise margin and input metastability point of view [5]. It can also be observed from Figure 4.6 that Si based inverters show much higher undefined region compared to its strained-Si counterpart for a particular channel length. In addition, it has been noted that there is very little dependence of logical undefined region on channel length in strained-Si inverters unlike Si inverters. This property, in essence, provides an extra degree of freedom to the designers to engineer a required noise margin by adjusting the width of the devices. The advantage with the strained-Si devices is that the DC transfer characteristics can move right or left with the same W_n/W_p ratio, by changing the amount of strain in n- or p-MOSFET transistor as the mobility enhancement of both these MOSFET for

different amount of strain are different. This will give enough flexibility for the digital designers to adjust the noise margins for saving from metastability and for the signal processing engineers for better data coupling.

4.2.2 Transient Characteristics

In analysing the transient characteristics of strained and standard CMOS circuits, the main attributes of interests are delay, slew rate and power consumption. The analyses of inverter based on strained and conventional Si technologies are presented in the following sub-sections.

4.2.2.1 Delay

The transient characteristics of the CMOS inverters built using strained-Si and Si technologies are shown in Figure 4.7 where it can be seen that there is an improvement in speed for strained-Si over conventional Si inverters. The results suggest that strained-Si CMOS devices are a better alternative to a conventional Si CMOS for high speed applications.

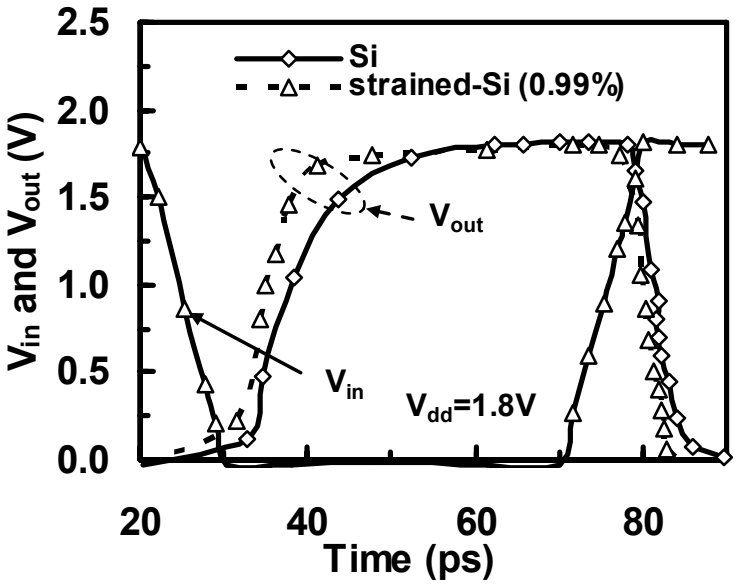


Fig. 4.7 Comparison of transient characteristics of CMOS inverters realised using strained-Si and Si technology.

The propagation delay for the inverters built using strained-Si and Si for different technology nodes ranging from 130 nm to 270nm are plotted in Figure 4.8. It should be noted that for all the technology generations strained-Si excels Si

technology in terms of speed. For all the technology generations considered; there is a speed improvement of around 26%. It can be concluded that for deeply cascaded circuits where the delay becomes additive [19], strained-Si will be much faster compared to circuits using the conventional Si technology that is, 26% faster. The dotted lines in Figure 4.8 indicate that strained-Si technology at 270nm exhibits same delay as that of standard Si technology at 130nm. Hence strained-Si is a better technology to combat device scaling issues.

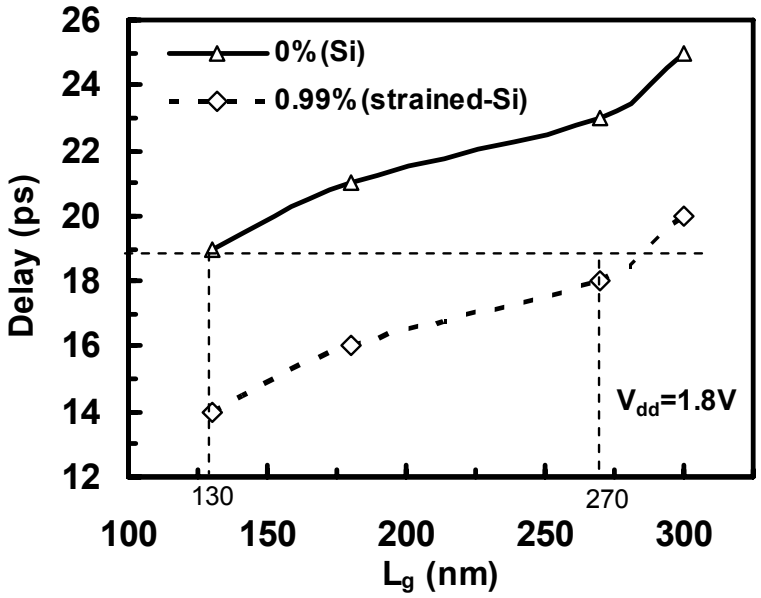


Fig. 4.8 Variation of gate delay against variation of gate length.

The results from the analysis of the inverter gate indicate that strained-Si is a potential candidate for deeply cascaded circuit applications with similar delay and can drive more inputs compared to the conventional Si CMOS device. Figure 4.9 show that the strained-Si CMOS inverter has a considerably smaller delay for the same fan-out compared to a conventional Si CMOS inverter at different gate lengths.

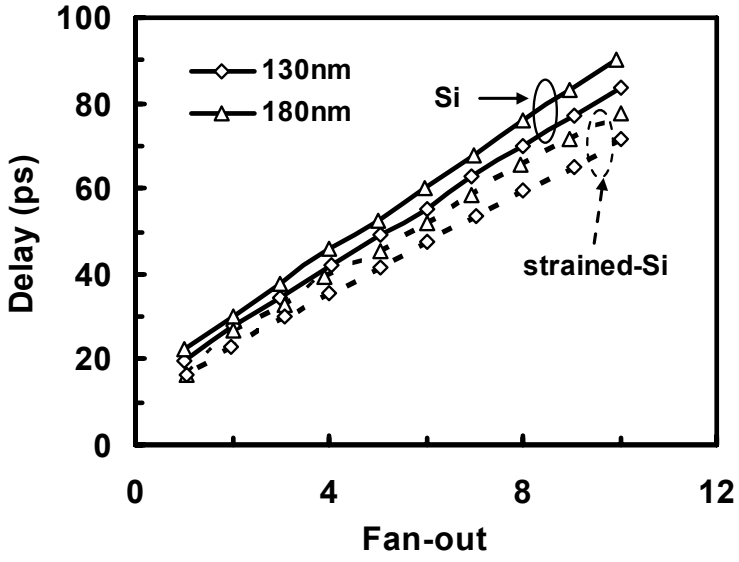


Fig. 4.9 Delay (ps) as a function of fan-out for CMOS inverters realised using strained-Si and Si technologies.

Even though the results show an improvement in speed for strained-Si technology, the static power dissipation may be slightly higher in strained-Si inverters for the same supply voltage due to higher transistor off state leakage current (I_{off}) in strained-Si (depending on the percentage of Ge in the virtual substrate) than the conventional Si CMOS technology which is discussed in Section 3.3.2 of Chapter 3. However, a comparable performance for strained Si with standard Si device can be achieved in terms of the off state power but by maintaining high speed, by lowering supply voltage and without scaling down the dimensions.

Figure 4.10 shows the variation of propagation delay for strained-Si and Si inverters for a supply voltage variation of 0.5V to 3.3V; the inverters were implemented using 130nm technology. Figure 4.10 also shows the propagation delays for strained-Si inverters with different amounts of strain in the channel. It should be noted that a device with 0% strain is basically the conventional Si device. It is interesting to note that the propagation delay of a strained-Si inverter with 0.99% strain in the channel consistently shows a lower delay (faster device) at all supply voltages. For all the simulations, the width of the device is considered to be 1μm. The analysis shows that the increase of delay with supply voltage reduction is

more severe for Si inverters. The strained-Si has a lower threshold voltage compared to that of the Si device due to the presence of strain and smaller band gap [20, 21]. The smaller delay exhibited by strained-Si device is accounted for by the enhanced mobility of the carriers, which depend on the amount of strain, induced in the channel [8, 22].

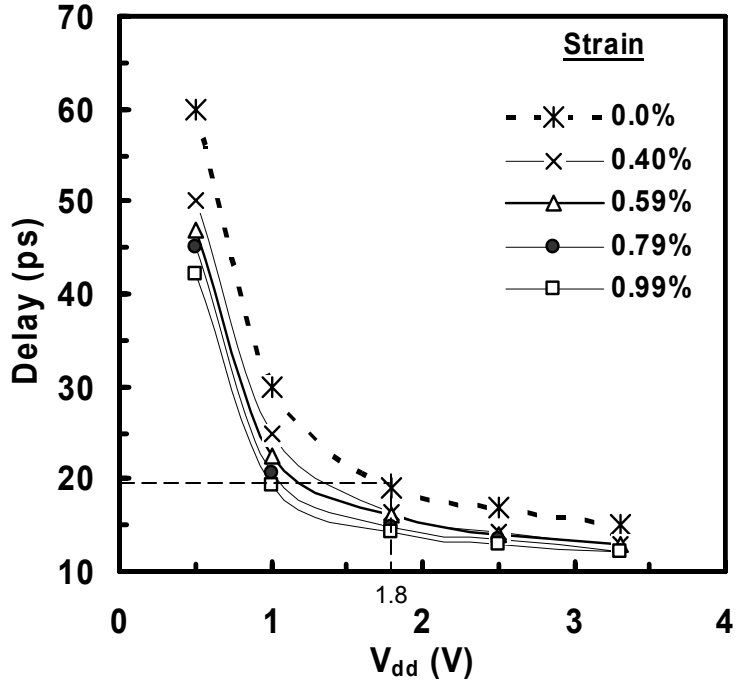


Fig. 4.10 Propagation delay vs Supply voltage variation of inverter realized using strained-Si and conventional Si CMOS devices on a 130nm technology node.

Compared to the strained-Si inverter with 0.99% strain in the channel, the delay for the Si inverter is seen to be higher (slower device) by 33%. It can also be seen from Figure 4.10 that the propagation delay of strained-Si inverter with a strain of 0.99% operated at a supply voltage of 1V is similar to that of the Si CMOS inverter operated at 1.8V (dotted lines in the figure). Thus a similar speed performance can be obtained for strained-Si devices compared to the Si one at a much lower supply voltage, resulting in significant reduction of dynamic power dissipation. This particular property of strained-Si circuits make them more attractive for ultra low-voltage (and consequently, ultra low-power) operation and thus, strained-Si circuits have greater potential than the Si-based circuits to meet the ITRS roadmap leading to ultra-low supply voltage era. Consequently, at ultra-low

supply voltages, the Si-based circuits exhibit low-power performance compromising its speed. These results demonstrate that the strained-Si based circuits can attain higher speed performance without compromising power even if the supply voltage reduction continues. However, reducing the supply voltage to achieve low power will degrade the noise margins [1, 23]. The variation of noise margin of inverters against the amount of stress applied in the channel of transistors was discussed in section 4.2.1.

4.2.2.2 Slew Rate

Another important metric which decides the performance of a circuit is the slew rate which is calculated from the DC and Transient charecteristics of a circuit. The slew rate of a device is the rate of change of its output voltage [9]. The change of output can be from a low voltage state to a high voltage state or vice versa. It is usually calculated as the time the device takes to switch from 10% to 90% of the final value in calculating the slew rate Equation 4.5 [24] is used:

$$\frac{dv}{dt} = \frac{(V_{OH} - V_{OL})}{t_r} \qquad \text{.....(4.5)}$$

Where, V_{OH} = output voltage at input voltage of V_{IL}.

V_{OL} = output voltage at input voltage of V_{IH}.

t_r = rise time of the signal (or fall time).

The slew rate of inverters built from strained-Si and Si are compared in Figure 4.11. The difference in slew rate is due to the change in the 'ON' resistance of the transistors [9]. Slew rate is compared for different supply voltages of 1V and 1.8V and when the amount of strain in the channel varies from 0% to 0.99%. It can be seen from the figure that the slew rate which reflects the slope of the DC and transient characteristics of a circuit increases with increase in the amount of strain which indicates the improved performance of strained-Si circuits compared to the Si circuits.

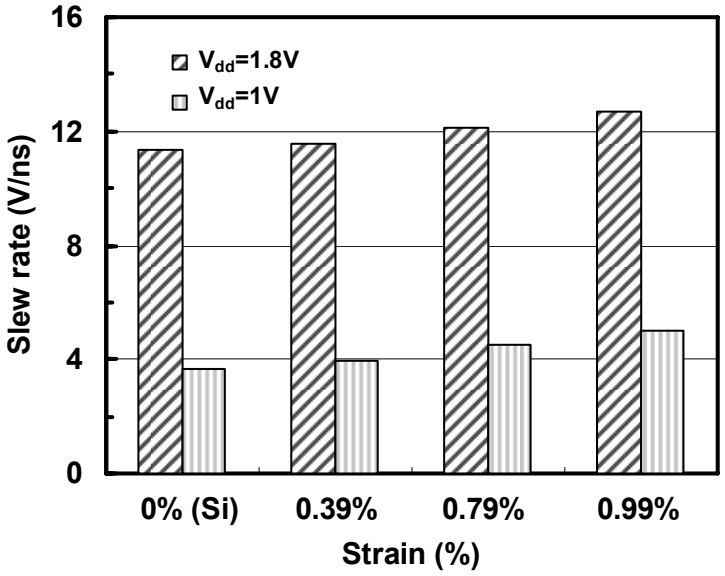


Fig. 4.11 Comparison of Slew rate of strained-Si and Si inverters

4.2.2.3 Ring Oscillator

Ring oscillators are often used as prototype circuits to test new semiconductor processes, because they are simple and easy to design. They are ideal test structures to optimize design parameters and layout rules for a new process. A Ring oscillator circuit with odd number of stages is shown Figure 4.12.

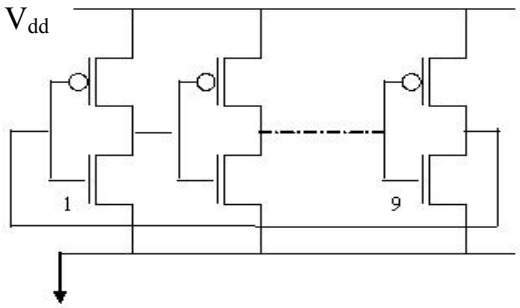


Fig. 4.12 The ring oscillator circuit with odd number of stages.

Ring oscillators using both strained-Si and conventional Si inverters were simulated and their frequency of operation is compared in Figure 4.13. It can be seen from the figure that the speed of operation of strained-Si based CMOS ring oscillators has been enhanced in comparison to the conventional Si based CMOS ring oscillators for the technology generations considered. A remarkable enhancement of 17 % in the frequency of operation is achieved for a 130 nm gate length strained-Si CMOS ring oscillator compared to conventional Si based CMOS counterpart. A very good speed improvement is expected by using strained-Si based circuits as deeply cascaded digital circuits provide bigger load capacities.

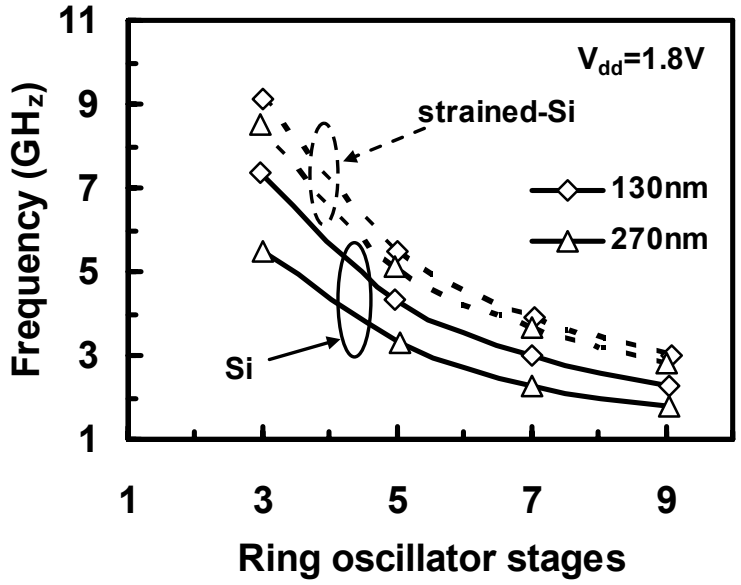


Fig. 4.13 Ring Oscillator frequency as a function of number of stages used.

4.2.3 Energy Efficieny

The results discussed in the previous sections shows that the strained-Si technology outclasses the conventional Si technology in terms of its advantage to build faster circuits and its usage in low-power circuits by scaling the supply voltage. In this section the energy and power related issues are discussed.

Figure 4.14 shows energy efficiency of an inverter realised using strained and conventional Si technology for different supply voltages. Minimum energy point (MEP) is defined as the supply voltage voltage point where the active energy

and leakage energy cross each other and is calculated as shown in the inset of Figure 4.12 [1, 25, 26]. It is apparent from the plot that the MEP shifts towards right with increasing induced strain in the channel.

MEP is found to vary from 0.57V to 0.7V depending on the amount of induced strain. The MEP for Si inverters for the same output load will fall below its threshold voltage and hence, these will not be functional at these supply voltages. This is due to the fact that the leakage current for a Si device at similar drain to source voltage is approximately an order of magnitude lower than that of a strained-Si device as shown in Figure 4.16 of Chapter 3.

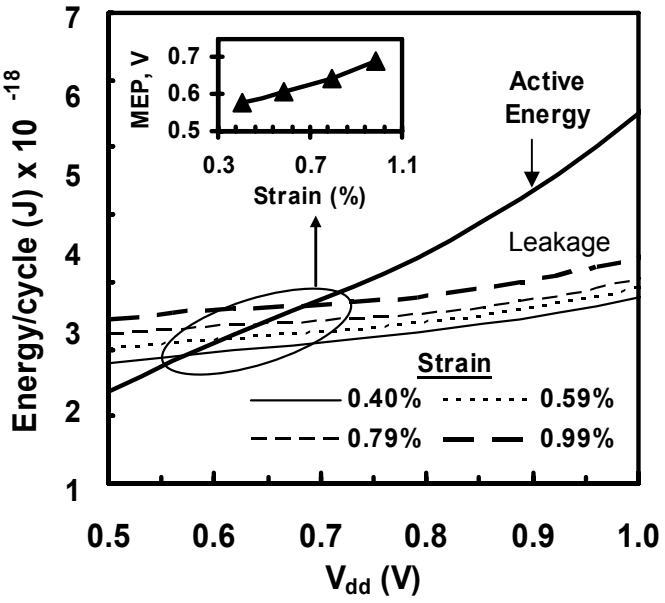


Fig. 4.14 The variation of Energy/Cycle with supply voltage for strained-Si inverters with varying amount of strain in the device channel.

If the strained-Si inverters with 0.99% strain in the channel are operated at a V_{dd} of 0.7V (which is the MEP for 0.99% strain) then the delay is found to be 32ps (Figure 4.10). However, it is found to be 40ps for the inverters with 0.40% strain in the channel at the same V_{dd}. This implies an improvement in speed of 20% for the inverter with 0.99% strain compared to those having 0.40% strain in the channel. If the operation of the inverters is considered from the perspective of minimum energy point (MEP) for 0.40% strain in the channel, then the delays of 47ps and 37ps are observed for the inverters with 0.40% and 0.99% of strain, respectively. This implies a 21% increase in speed for the inverters with 0.99% strain compared to the inverters

with 0.40% strain in the channel. However, this speed improvement will compromise the minimum energy criteria of 0.99% strain.

The important point here is that depending on the circuit requirement (whether faster operation or lower power) strain can be engineered to satisfy the required criteria. Although the natural tendency is to increase speed by using more strain, when all the performance parameters such as noise immunity, power dissipation and speed are considered, the result shows that 0.99% (more leakage) strain or 0.40% (slower operation) strain could be an optimal solution. The optimal strain can be be chosen following the need of a desired set of circuit performance parameters. This added flexibility of strained-Si process is very much beneficial as circuits with different properties can be obtained under the same fabrication process.

4.2.4 Power Analysis

Power consumption is a main concern when devices are scaled to achieve speed with more functionality per unit area. Supply voltage reduction has been the preferred technique for power reduction. However, when the devices are approaching atomic level dimensions, the leakage current is increasing dramatically which subsequently leads to increased static power dissipation [2, 27, 28]. To analyse and compare the power dissipation of the two technologies; strained-Si and conventional Si, devices with in the range of 45nm to 100nm technology node are chosen for analysis. The parameters required for the simulations are taken from MASTAR [29] and ITRS [3].

Transient simulations of the strained-Si and Si inverters are performed for different technology nodes at a supply voltage of 1.2V; the rise time and fall time of the input pulse are taken as 10ps. Here, for the analysis the simulations are performed for strained-Si with 0.99% strain. The ratio of widths of p- MOS to n-MOS transistors is 2μm:1μm in the CMOS inverter to achieve nearly symmetrical DC characteristics. However, it should be noted that the enhancement of mobility of holes and electrons due to different amount of strain [10] are different and hence there is an added flexibility for the designers in choosing the width of the devices.

Figure 4.15 shows the plot of gate delay of Si and strained-Si CMOS inverter against technology generations ranging from 45nm to 100nm. It shows an increase of delay for both types of devices. However, it is more severe in Si devices (slow) at large device sizes. It can be seen from the figure that the propagation delays for a 65nm Si CMOS inverter is approximately equivalent to that of strained-Si which is two generations older (dotted lines in Figure 4.15) obviating the adverse effects of scaling to some extent. These results are attributed to the fundamental difference of Si and strained-Si which were explained in the preceding chapters.

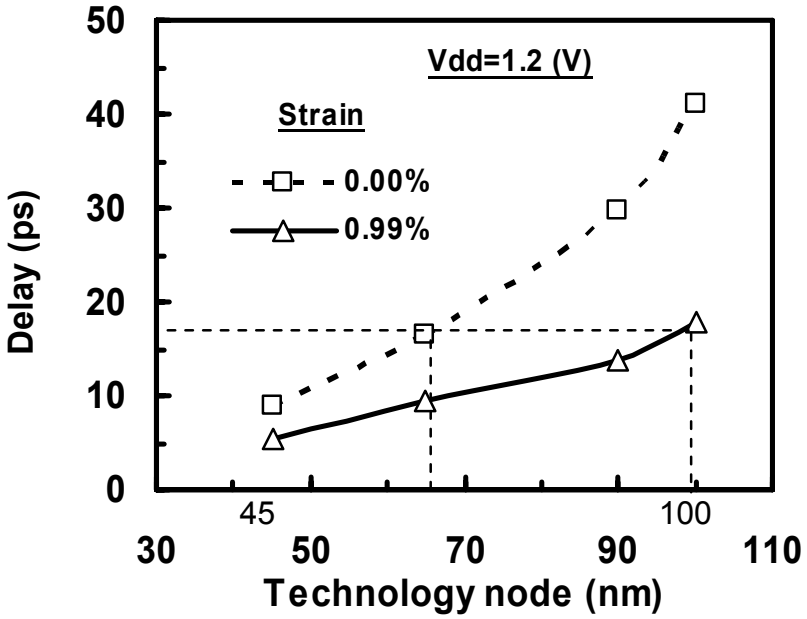


Fig. 4.15 Gate delay of strained-Si and Si inverters for different technology nodes.

It has also been observed that when the device dimensions are scaled down strained-Si based inverters consistently show a smaller delay compared to its Si based counterpart. It was also apparent from the results that the difference in delays between strained and conventional Si CMOS inverter becomes slightly lower with the decrease of channel length. From the above results shown in Figure 4.15, strained-Si inverters have almost 2x speed advantage over its Si counterpart and this speed advantage can be exploited to reduce power dissipation by supply voltage scaling for a comparable speed performance. However, the leakage current and

subsequently the static power dissipation of strained-Si are higher than that of Si due to the inherent transport properties of strained-Si.

Figure 4.16 shows the static power dissipation of inverters realized with strained-Si and Si technology for a supply voltage of 1.2V. Different device dimensions are considered here for the power calculation.

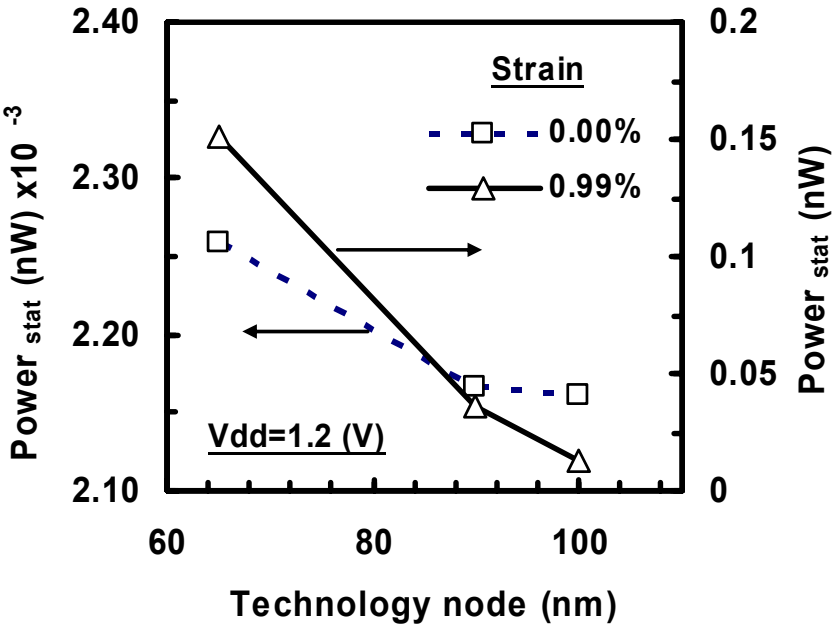


Fig. 4.16 Static power dissipation of strained-Si and Si CMOS inverters for different technology generations at a supply voltage of 1.2V.

It can be seen from the figure that the static power dissipation of a Si inverter is almost an order of magnitude lower compared to a similar strained-Si CMOS inverter and is increasing for all technology generations. This is attributed to the fact that strained-Si has a considerably smaller band gap than that of Si material which leads to higher subthreshold current [10, 14]. The static power increases with increase in the amount of strain applied. This can be seen as a compromise to the performance of strained-Si. However, this speed advantage can be exploited to reduce power dissipation by supply voltage reduction for a comparable speed performance as discussed in section 4.2.2.1 and shown in Figure 4.10. It has been observed from the results that even with supply voltage reduction, strained-Si outperforms Si based circuits.

Figure 4.17 shows the variation of the dynamic power [23] of inverters realised with strained-Si and Si in different technology generations. The figure shows that the dynamic power of a strained-Si CMOS inverter is higher than that of Si for all the technology generations.

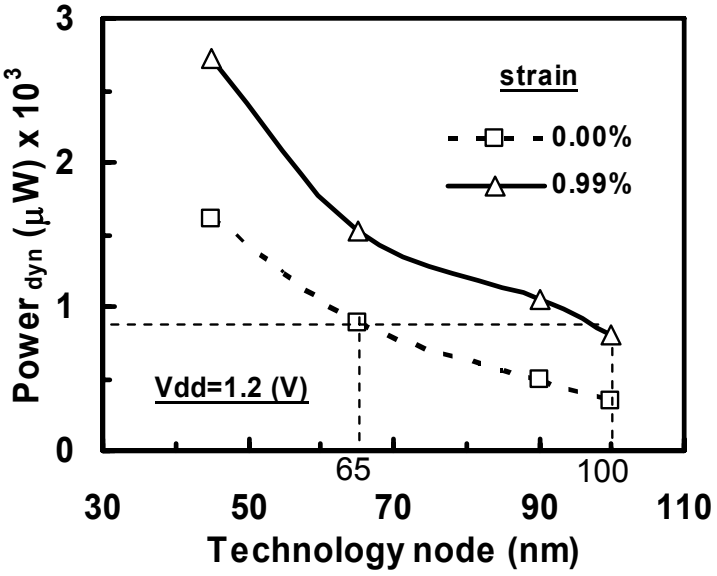


Fig. 4.17 Dynamic power of Si CMOS inverter is compared against to that of strained-Si which is strained by an amount of 0.99% for different technology generations.

The reason for the increased dynamic power dissipation is obvious due to the fact that the drain current of strained-Si transistors is larger compared to Si inverter for the same supply voltage. However, reducing the supply voltage has a quadratic effect on the dynamic power [30]. Figure 4.18 shows the dynamic power variation of strained-Si and Si CMOS inverters against the reduction of supply voltage for 65nm technology node. It can be seen that for the same dynamic power (marked as dotted lines), the supply voltage for a strained-Si CMOS inverter is almost 1V; whereas the supply voltage for the conventional Si CMOS inverter is 1.2V. However, the reduction of supply voltage will degrade the performance.

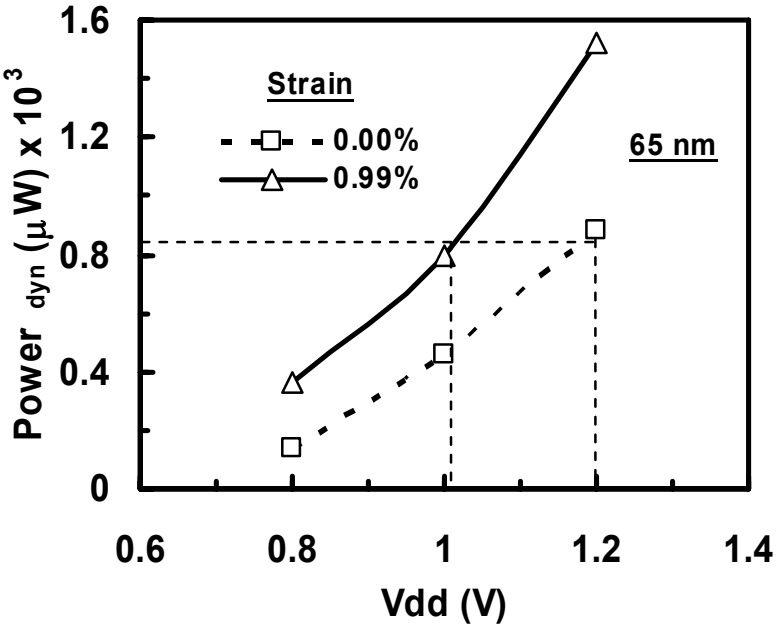


Fig. 4.18 Dynamic power of strained-Si against that of Si CMOS inverter when the supply voltage is scaled.

The variation of gate delay against the supply voltage reduction was shown in Figure 4.10 where, strained-Si circuit is found to be much faster than Si. However, the reduction of supply voltage poses a challenge on reducing the noise margin. It has already observed from our previously published results that noise margin conditions are better for strained-Si circuits than that of Si circuits.

The energy delay product (EDP) is a metric which combines performance and energy for a process technology [31]. It can be seen from Figure 4.19 that the EDP of strained-Si is almost 100% less than that of corresponding Si at the 65nm technology node for the same supply voltage of 1.2V. Furthermore for the same supply voltage strained-Si shows almost 100% improvements in terms of speed. It can also be noted that the EDP of Si is getting worse at higher technology nodes and is due to the fact that the speed of the Si circuit is degraded faster than that of strained-Si circuits for higher technology nodes. The minimum energy point (MEP) for strained-Si circuits with different amount of strain has been observed previously which proved the added flexibility of strained-Si circuits in the design arena.

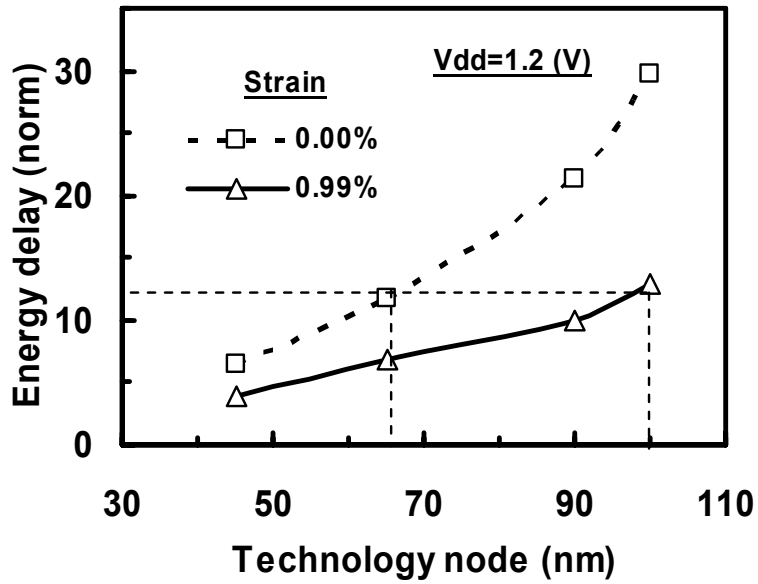


Fig. 4.19 Energy Delay Product (EDP) of strained-Si and Si CMOS inverters for different technology generations.

4.2.5 Strained-Si and Si devices designed for same leakage

Even though the results discussed in Sections 4.2.1 to 4.2.4 indicate that strained-Si technology is a better candidate for high-speed and low-power circuit applications there is a need to reduce the power supply voltage for strained-Si circuits to consume less power. When strain is applied in the channel, the leakage current is found to increase due to the inherent property of the lower band-gap in the strained-Si material. Figure 4.20 shows the decrease of leakage current flowing through the strained-Si and Si transistors when the drain/source voltage is reduced.

From Figure 4.20, it is seen that the leakage current of a strained-Si device is higher than that of Si by almost an order of magnitude. For continuously operated circuits, the increased leakage current due to reduction in threshold voltages is not significant. However, circuit components, which are idle during the operation of a circuit, suffer from the effects leakage currents resulting in an increase in static power dissipation.

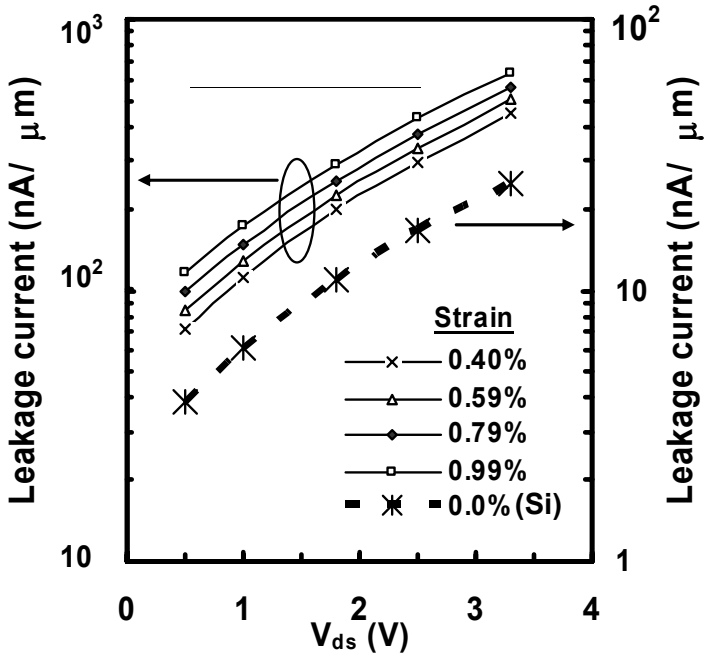


Fig. 4.20 Leakage current Vs drain source voltage for different amount of strain along the device channel.

Equation 4.6 shows the relation between subthreshold current and device and process parameters [27, 30].

$$I_{sub} = u_0 C_{ox} \frac{W_{eff}}{L_{eff}} V_T{}^2 e^{1.8} \exp\left(\frac{V_{gs} - V_{th}}{nV_T}\right) \cdot \left(1 - \exp\left(\frac{-V_{ds}}{V_T}\right)\right) \ldots\ldots (4.6)$$

From Equation 4.6 it can be seen that a reduction in I_{sub} can be achieved at device level by:

- Reducing the V_{gs} and V_{ds} in the device.
- Decreasing the width of the device or fabricating devices using a larger technology node (L_{eff}).
- Reduce oxide capacitance C_{ox} by increasing t_{ox} (this is determined by the foundry).
- Increasing the threshold voltage V_{th} by increasing the doping which can be altered by varying the diffusion time, energy and doping density.

To check whether a strained-Si based circuit is still faster when designed by keeping the same leakage current as that of the Si based inverter, strained-Si transistors were designed by changing the process recipes to leak the same current as that of Si transistor with same dimensions and at the same supply voltage.

Figure 4.21 shows the comparison of CMOS inverter delay made from strained-Si devices which is designed to give the same off current as that of Si devices (with similar dimensions). 0.99% strain is considered here for the analysis. It can be seen from the Figure 4.21 that the strained-Si CMOS inverter exhibits a lower delay than that of the corresponding Si device in all the technology generations from 65nm to 100nm and strained-Si inverter is giving an average speed improvement of around 16% which again demonstrates that strained-Si is a better candidate for low-power high speed applications. Therefore even with the same supply voltage conditions, by changing process recipes, still strained-Si is found to be the best for high-speed low-power circuits. This is an advantage for circuits which drive deeply cascaded circuits which will still gives a performance improvement. Besides these process changes to adapt the circuits for certain applications, design methods can be used to reduce power consumption to save severe costs which incur when the processes are changed [32]. Stacking of transistors [32, 33] is such a method that will be discussed in the next section.

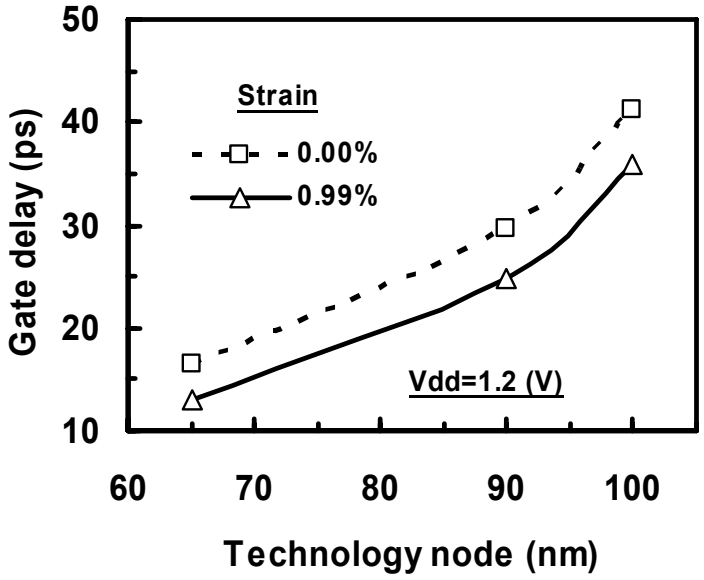


Fig. 4.21 Speed of Si and strained-Si CMOS inverters designed for same static power is compared for different technology generations.

4.2.6 Device Stacking

The method of device stacking is a proven technique used by designers to reduce the subthreshold leakage current in the stand-by mode of operation of a circuit subsequently reducing the static power consumption [33, 34]. The method of transistor stacking is shown in Figure 4.22 where n-MOS transistors (half of the normal width) are stacked in the pull-down path of a CMOS inverter. Unfortunately, device stacking affects the performance adversely [33-36]; however, it will be demonstrated that a stacked strained-Si inverter still outperforms a Si inverter, whilst consuming the same amount of power. In this section stacked inverters are built using strained-Si technology and compared with Si stacked and non-stacked CMOS inverters. The whole analysis with stacking method is performed for devices with different supply voltage conditions and varying amount of strain.

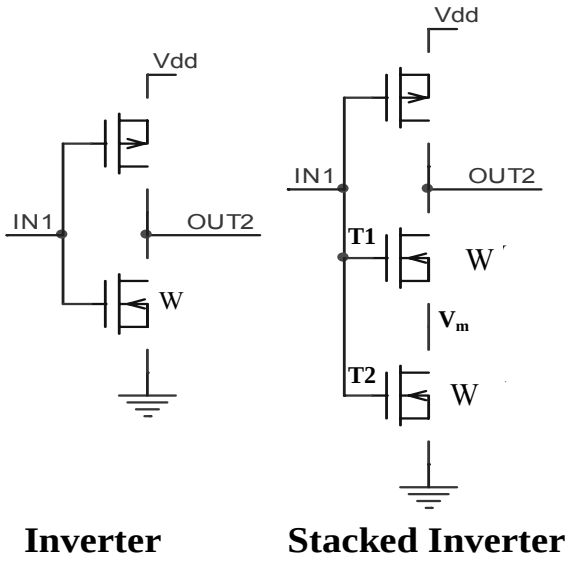


Fig. 4.22 Schematic of non-Stacked and stacked inverter

4.2.6.1 Stacking Theory

When devices are stacked as shown in Figure 4.22, during the 'OFF' condition of stacked the transistors the intermediate node voltage (marked as Vm in Figure 4.22) is forced to go to a value higher than zero [36].

When T1 and T2 (stacked transistors in Figure 4.22) are turned off, the voltage, V_m, at the intermediate node is positive due to a small drain current thus;

- $V_{gs1} < 0$, thus reduce the subthreshold current substantially.
- $V_{bs1} < 0$, increase V_{th1} (larger body effect) and thus reducing the subthreshold leakage current
- When V_{ds1} decreases, V_{th1} increases.
- V_{ds2} decreases; increasing V_{th2}.

Due to the increase in V_{th1} and V_{th2} the subthreshold current reduces dramatically leading to reduced static power.

Figure 4.23 shows the variation of static power consumption of stacked and non-stacked inverters against strain. The inverter supply voltage is 1.1V and the device parameters used are for the 65nm technology node for Low Standby Power (LSTP) specification provided by ITRS [3] and MASTAR [29].

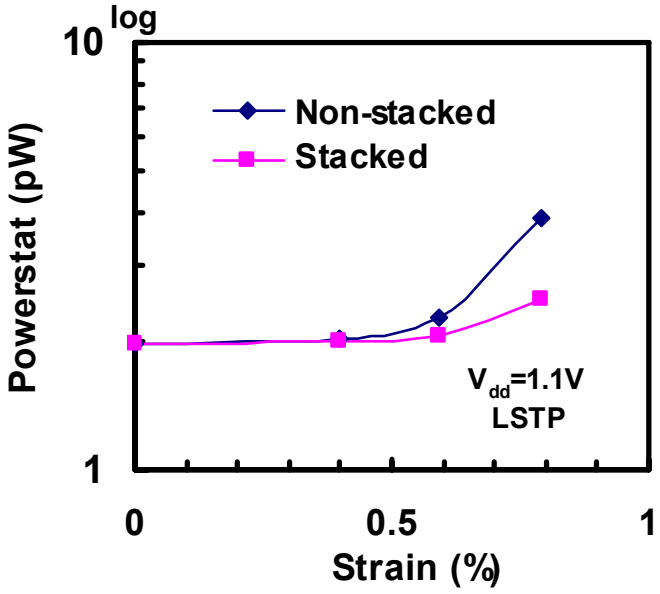


Fig. 4.23 Comparison of static power dissipation of stacked and non-stacked CMOS inverters as a function of strain.

The amount of strain in the channel is varied from 0% (Si) to 0.99%. It can be seen from the figure that up until 0.6% strain, the static power

consumption of stacked strained-Si inverter (strain >0) is almost the same as that of Si (strain=0) stacked inverter although it is increasing. At 0% strain, the leakage current of both non-stacked and stacked inverters are in the order of nano-amperes and hence the difference in the power dissipation is not noticeable. However, at larger amounts of strain the power dissipation of non-stacked inverters increases considerably. Since the threshold voltage of the stacked devices increases [33], the speed of the circuit is adversely affected. Hence, the method of device stacking is usually applied in circuit paths that are not critical.

A plot of the variation of gate delay of stacked and non-stacked inverter configurations against the percentage strain in the device channel is shown in Figure 4.24.

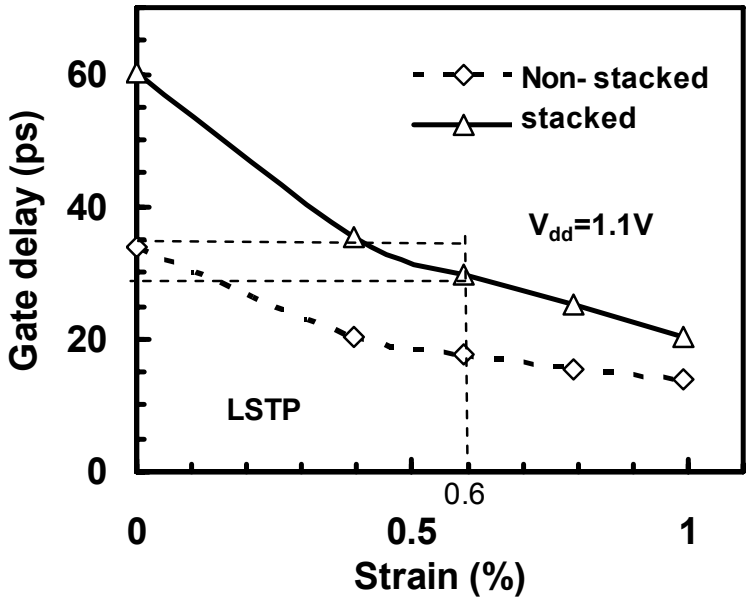


Fig. 4.24 Comparison of the delay of stacked and non-stacked inverters as a function of percentage strain in the device channel.

The figure shows an increase in gate delay for stacked inverters making them slower compared to the non-stacked inverters. At lower amounts of strain the delay increase is severe in Si devices (slow) (0 % strain). It has already been observed from Figure 4.23 that the static power dissipation of stacked inverter with 0.6 %

strain and a supply voltage of 1.1V is almost similar to non-stacked Si inverter. However, it is interesting to note from Figure 4.24 that the propagation delay of stacked inverter with 0.6% strain on a 65nm technology node with a supply voltage of 1.1V is almost 21% faster than a Si CMOS non-stacked inverter (meant to be faster) which demonstrates that strained-Si maintains improved performance than Si circuit even when design techniques to reduce power are applied.

From these results it can be concluded that strain with 0.6% can be considered as an optimum amount of strain for high-speed and low-power circuits as the static power is seen to be higher for strain with 0.99% and this is attributed to its smaller band-gap. However, due to the degraded performance of stacked inverters compared to non-stacked inverters, it is possible to apply stacking in non-critical path (paths which are not determining the operational speed of a circuit) where the transistors are mostly in stand-by mode operation to reduce the total power consumption.

Figure 4.25 compares the gate delay of stacked strained-Si (strain=0.6%) with non-stacked Si inverter for a range of supply voltages. It can be seen from the figure that for all the supply voltage conditions strained-Si stacked inverter is seen to be faster than that of the equivalent Si inverter with no stacking.

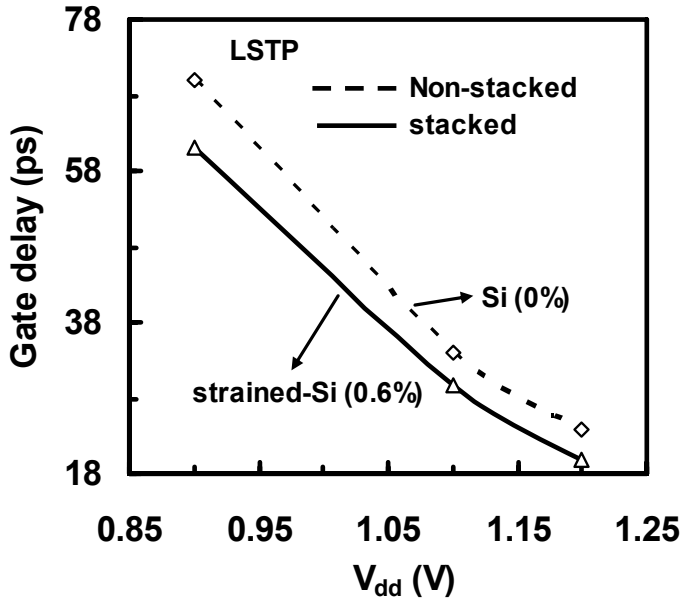


Fig. 4.25 Comparison of strained-Si stacked inverter with a non stacked Si inverter against variation of supply voltage.

4.6.2.2 Stack Factor

The Stack Factor is defined as the ratio of leakage current in one 'OFF' device to the leakage current in a stack of two 'OFF' devices [33]. In Figure 4.26 the stack factor is plotted against the amount of strain in the channel of transistors used to build the inverter. It can be seen from the figure that, until the strain reaches 0.6% the stack factor is almost the same as that of the Si inverter and thereafter it increases (degrades).

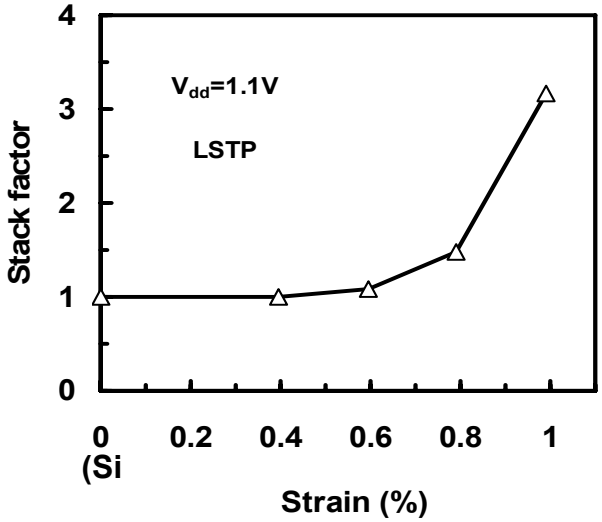


Fig. 4.26 Stack factor of CMOS inverter based on 65nm technology node designed for Low standby power (LSTP) operation.

4.3 CMOS Inverter Amplifier

It has been established in previous sections that the strained-Si based digital circuit offers several advantages, and is suitable for high-speed digital circuits that require less static power dissipation. In this section, the analysis is performed to identify the advantages of using strained-Si technology in analogue circuit applications. The results are demonstrated based on the frequency analysis of a CMOS inverting amplifier. The devices considered here are strained-Si and Si transistors built to 90nm and 100nm technology specifications. A strain of 0.99% in the device channel is considered for the analysis.

Figure 4.27 shows the gain-frequency plot of strained-Si and Si CMOS inverting amplifier. The negative sign for the gain is due to the inherent property of

common source amplifiers (180^0 phase shift) [37]. The higher gain of the strained-Si amplifier compared to that of Si is due its higher transconductance [10, 14]. It should also be noted from the figure that the strained-Si circuit amplifies a larger range of frequencies with higher gain compared to the Si one and hence gives larger bandwidth. The horizontal lines in Figure 4.27 (dotted for Si and solid for strained-Si) indicate the 3dB point of the amplifiers to calculate the cut off frequency of the amplifier [37]; the arrow shows the variation of higher cut-off frequency of the inverting amplifier when the strain is increased from 0% (Si) to 0.99% (strained-Si) in the 90nm transistors. The higher cut off frequency for the strained-Si and Si CMOS amplifiers (90nm) is found to be 200 MHz and 500 KHz respectively.

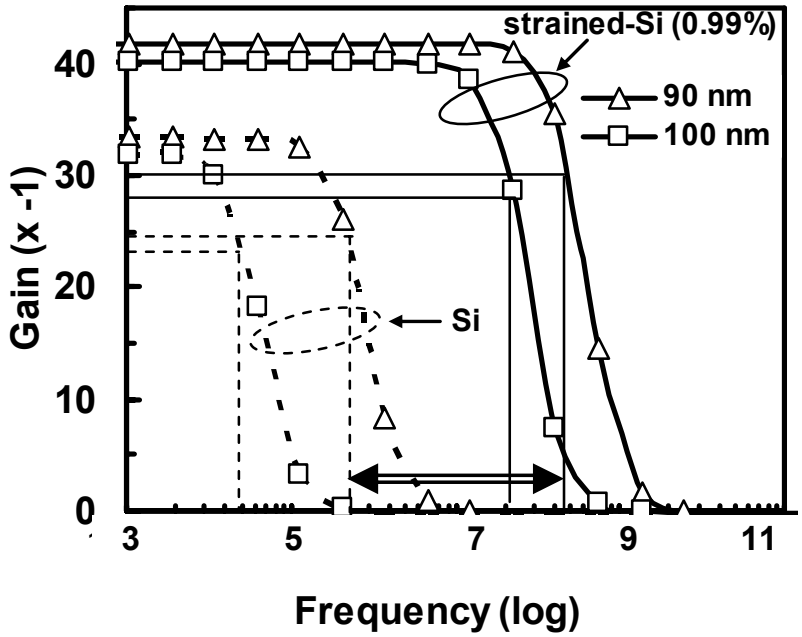


Fig. 4.27 Gain Vs High frequency cut-off for a strained-Si and Si CMOS inverting amplifier.

Figure 4.28 shows the bandwidth variation of the CMOS inverter amplifiers realised with strained-Si and Si technology. It can be seen that when the technology node is lowered, the bandwidth is increasing. In addition, for larger amounts of strain, the bandwidth is seen to improve which makes strained-Si technology a better alternative to Si technology for analog applications.

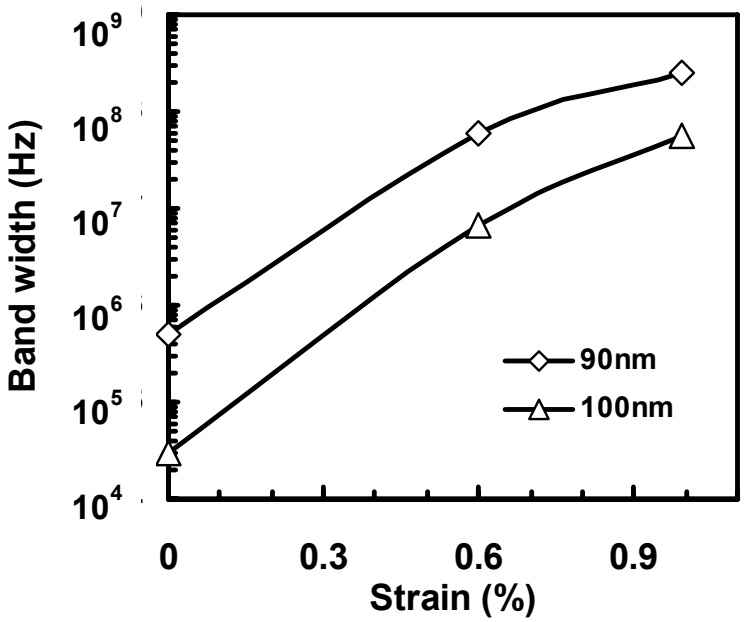


Fig. 4.28 Variation of bandwidth against the amount of strain in the device channel.

4.4 Summary

In this chapter we have explored the potential of strained-Si for low-power digital circuits for example, microprocessors, portable electronics, communication equipments etc and its flexibility in designing amplifiers with specific bandwidths for different applications. The results shows that it is possible to achieve significantly improved low-power performance from strained-Si based circuits than Si, since it allows significant window of voltage reduction without sacrificing much delay which is due to the inherently low threshold voltage and high carrier mobility of strained-Si devices. Thus, under the same performance specification, the strained-Si based circuits can achieve better power reduction compared to its Si-based counterpart by voltage reduction.

However, the supply voltage scaling adversely affects the noise margins of the circuit which leads to for example, metastability failure. Hence, emphasis has also been given to the evaluation of the noise characteristics, low-power performance and delay characteristics under different channel straining and operating conditions. It is seen that the strained-Si inverter is still better compared

to that of the conventional Si inverters. However in the sub-1V supply region, the noise performance is relatively independent of the amount of induced strain. The results show great promise for strained-Si technology in digital applications which require high throughput and low static power dissipation without degrading the noise characteristics.

The Minimum Energy Point (MEP) of strained-Si also demonstrates its ability to operate at supply voltages near to its threshold voltage without incurring much leakage current while at the same time reducing the dynamic power significantly. The additional advantage of strained-Si based circuits is that the strain can be engineered to satisfy different types of circuit requirements under the same process flow. These facts make strained-Si circuits an ideal candidate for futuristic ultra low-power digital circuit design. A combined approach from design and technology perspective to reduce the power consumption by using stacking of transistors with strained-Si technology also demonstrates its potential dominance over the conventional Si technology.

The frequency analysis of a CMOS inverting amplifier; realised with strained-Si technology, shows an improved performance compared to a Si based amplifier; almost 2x bandwidth is achieved with strained-Si technology compared with the Si counterpart with similar dimensions. These results coupled with the availability of higher supply voltage reduction window also make strained-Si based circuits an ideal candidate for analog circuit design.

The Table 4.1 below summarizes the comparison of circuit atttributes of strained-Si and Si technologies.

Characteristics	Si	Strained-Si	Advantages
Noise Margins (mV)	NML=350 NMH=1100	For 1% strain NML=450 NMH=1250	Better Noise immunity
Undefined region (V)	0.28	0.35	Less prone to metastability
Speed (ps)	19	14	Faster circuits
Fan-out of 9- delay (ps)	85	65	Deeply cascaded Circuits
Ring oscillator Frequency (9 stages)	1GHz	3GHz	Faster circuits
Static power		Higher	Power consuming circuits (disadvantage)
Static power after stacking (0.6% strain)	1.8pW	2pW	Similar power consumption
Speed after stacking	60ps	20ps	Faster critical paths
CMOS inverting amplifier frequency	500KHz	200MHz	Remarkable improvement for strained-Si
Bandwidth of CMOS inverting amplifier	10^4	10^8	Suitable for communication applications.

Table. 4.1 The summary of the results.

References

[1] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-Power CMOS Digital Design", *IEEE Journal of Solid-State Circuits*, vol. 27, pp. 473-484, 1992.

[2] K. Roy and S. Prasad, *Low-power CMOS VLSI circuit design*. New York: Wiley, 2000.

[3] International Technology Roadmap for Semiconductors, "Process Integration, Devices and Structures", 2007 Ed.

[4] C. Fiegna, H. Iwai, T. Wada, and M. Saito, "Scaling the MOS Transistor below 0.1μm: Methodology, Device Structures, and Technology Requirements", *IEEE Transactions on Electron Devices*, vol. 41, pp. 941, 1994.

[5] J. F. Wakerly, *Digital design: principles and practices*. New Jersey: Prentice Hall, 1990.

[6] Z.-Y. Cheng, M. T. Currie, C. W. Leitz, G. Taraschi, et.al., "Electron Mobility Enhancement in Strained-Si n-MOSFETs Fabricated on SiGe-on-Insulator (SGOI) Substrates", *IEEE Electron Device Letters*, vol. 22, pp. 321-323, 2001.

[7] S. E. Thompson, G. Sun, Y. S. Choi, and T. A.-N. Nishida, T., "Uniaxial-process-induced strained-Si: extending the CMOS roadmap", *IEEE Transactions on Electron Devices,* vol. 53, pp. 1010-1020, 2006.

[8] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Fabrication and analysis of deep submicron strained-Si n-MOSFET's", *IEEE Transactions on Electron Devices,* vol. 47, pp. 1406-1415, 2000.

[9] N. H. E. Weste, D. Harris, and A. Banerjee, *Principles of CMOS VLSIDesign: A Circuit and Systems Perspective*. New Delhi: Pearson Education, 2007.

[10] C. K. Maiti, N. B. Chakrabarti, and S. K. Ray, *Strained Silicon Heterostructures: Materials and Devices*. London: Institution of Electrical Engineers, 2001.

[11] S. M. Sze, *VLSI technology*. New York: McGraw-Hill, 1983.

[12] D. A. Neamen, *Semiconductor Physics and Devices: Basic Principles*. Boston: McGraw-Hill, 2003.

[13] K. Seeger, *Semiconductor Physics*. New York: Springer-Verlag, 1973.

[14] S. Chattopadhyay, L. D. Driscoll, K. S. K. Kwa, S. H. Olsen, et.al., "Strained Si MOSFETs on relaxed SiGe platforms: performance and challenges". *Solid-State Electronics*, vol. 48, pp. 1407-1416, 2004.

[15] K. A. Jenkins and K. Rim, "Measurement of the Effect of Self-Heating in Strained-Silicon MOSFETs", *IEEE Electron Device Letters*, vol. 23, pp. 360-362, 2002.

[16] R. L. Peterson, "Stretching silicon: a uniaxial and biaxial strain generation process and the resulting mobility enhancement in silicon-on-insulator MOSFETs", Thesis (Ph. D.), Princeton University, 2006.

[17] Z. Cheng, J. Jung, M. L. Lee, A. J. Pitera, et.al., "Hole mobility enhancement in strained-Si/strained-SiGe heterostructure p-MOSFETs fabricated on SiGe-on-insulator (SGOI)", *Semiconductor Science and Technology.*, vol. 19, pp. L48-L51, 2004.

[18] L. J. Huang, J. O. Chu, S. A. Goma, C. P. D Emic, et.al., "Electron and Hole Mobility Enhancement in Strained SOI by Wafer Bonding", *IEEE Transactions on Electron Devices*, vol. 49, pp. 1566-1571, 2002.

[19] E. Sicard and S. Ben Dhia, *Advanced CMOS cell design*. New York: McGraw-Hill, 2007.

[20] P. M. Mooney, "Materials for Strained Silicon Devices", *IBM reserach report- Material Science*, pp. 1-11, 2002.

[21] V. Chan, K. Rim, M. Ieong, S. Yang, R. Malik, et.al., "Strain for CMOS Performance Improvement", *IEEE Custom Integrated Circuits Conference*, vol. 2005, pp. 667-674, 2005.

[22] S. H. Olsen, A. G. O Neill, P. Dobrosz, S. J. Bull, et.al., "Study of strain relaxation in Si/SiGe metal-oxide-semiconductor field-effect transistors", *Journal of Applied Physics*, vol. 97, pp. 114504, 2005.

[23] A. P. Chandrakasan and R. W. Brodersen, *Low-power CMOS design*. New York: IEEE Press, 1998.

[24] R. L. Boylestad and L. Nashelsky, *Electronic Devices and Circuit Theory*. New Jersey: Prentice-Hall, 1982.

[25] S. Narendra, V. De, S. Borkar, D. A. A.-A. Antoniadis, D.A., et al., "Full-chip subthreshold leakage power prediction and reduction techniques for sub 0.18μm CMOS", *IEEE Journal of Solid-State Circuits,* vol. 39, pp. 501-510, 2004.

[26] S. G. Narendra and A. P. Chandrakasan, *Leakage in Nanometer CMOS Technologies*. New York: Springer, 2006.

[27] S. M. Sze, *Physics of Semiconductor Devices*. New York: Wiley, 1981.

[28] J. P. Colinge, C. A. Colinge, and I. NetLibrary, *Physics of Semiconductor Devices.* Boston: Kluwer Academic Publishers, 2002.

[29] Model for assessment of CMOS technologies and roadmaps, *MASTAR,* ST Microelectronics*,* Advanced devices research group, 2000-2005

[30] P. R. Meer, A. Staveren, and A. H. M. Roermund, *Low-Power Deep Sub-micron CMOS Logic: Subthreshold Current Reduction*. London: Kluwer Academic, 2004.

[31] M. Elgebaly, "Energy-efficient design for deep sub-micron CMOS VLSIs", Thesis (Ph. D.), University of Waterloo, 2005.

[32] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits", *Proceedings of the IEEE*, vol. 91, pp. 305-327, 2003.

[33] S. Narendra, S. Borkar, V. De, D. A. Antoniadis, et.al., "Scaling of stack effect and its application for leakage reduction", *International Symposium on Low Power Electronics and Design*, 2001., 2001.

[34] S. Narendra, V. De, S. Borkar, D. A. Antoniadis, et.al., "Full-chip subthreshold leakage power prediction and reduction techniques for sub-0.18-/spl mu/m CMOS", *IEEE Journal of Solid-State Circuits,* vol. 39, pp. 501-510, 2004.

[35] J. P. Halter and F. N. Najm, "A gate-level leakage power reduction method for ultra-low-power CMOS circuits", *IEEE Custom Integrated Circuits Conference,* 1997.

[36] S. W. Amos and M. James, *Principles of Transistor Circuits: Introduction to the Design of Amplifiers, Receivers, and Digital Circuits*. Boston: Newnes, 2000.

CHAPTER 5

IMPACT OF PROCESS PARAMETER VARIABILITY ON STRAINED SILICON CIRCUIT PERFORMANCE

5.1 Introduction

The problems associated with scaling have been discussed in the literature review section of Chapter 2. The issues associated with scaling range from short channel effects in devices to the issue of power dissipation in CMOS circuits [1-3]. The use of strained-Si technology was suggested and discussed in Chapter 3 and 4 as an alternative to mitigate the issues with regards to the power consumption and speed associated with devices and circuits. However, variability (fluctuations in process parameters) in IC manufacturing is becoming a threat in the drive to scale down to deep submicron (DSM) chip dimensions [4]. Different sources of variability have already been explained in Chapter 2. The analysis of variability has become a very important tool, in the design of circuits with deep sub micron technologies, to predict the response variation very early in their design phase [5-7] due to process parameter spreads.

Furthermore, for several decades the output from semiconductor manufacturers has been high volume products with process optimisation being continued throughout the lifetime of the product to ensure a satisfactory yield. However, product lifetimes are continually shrinking to keep pace with market

demands. In addition, there is an increase in foundry business where product volumes are low; subsequently it is no longer feasible to optimise the process during the product lifetime resulting in an increase in parametric yield loss. Consequently there is an increasing need for semiconductor process versatility which can be adopted by a diverse customer base. Central to ensuring customer satisfaction of the manufactured product is an understanding of the effects of process variation on customer designs. By interfacing TCAD simulations [8, 9] with a SPICE model extraction strategy [10] and applying variations to parameters which are specific to the selected manufacturing process, designers are able to predict the effect of variation of process parameters thereby improving the yield or matching the manufactured devices, according to the characteristics to particular applications.

However, to study the impact of variability in the manufacture of transistors statistical methods seem to be an obvious choice as corner case methods, previously used do not seem to work for deep sub-micron manufacturing technologies [4]. Corner cases would prove to be an overdesign for new technologies which will incur severe cost and poor yield for the manufactured product. Table 5.1 summarizes the advantages of statistical design techniques over the conventional worst case design (corner case) approach. Previously the effects of process variations were accounted for by incorporating artificial distributions in SPICE models [10, 11].

Corner case design	**Statistical design**
Logic gates in a circuit have the same process corners	Different logic gates in the same circuit have different process corners
Suffers from overdesign	Realistic worst case extremes
Significant performance loss	Well designed circuits with no performance loss
Computationally efficient	Computationally inefficient

Table. 5.1 Corner case design Vs Statistical design techniques.

Monte Carlo Analysis is a widely used technique to analyse parameter variability to generate response distributions [12]. The simulation flow for the technique is shown in Figure 5.1. Although the technique is very accurate it is computationally inefficient particularly when a large number of variables is involved, for example, in analysing the effect of parameter variation in semiconductor processes; however, if appropriate, accuracy could be traded-off against computation times. A further disadvantage of this method is that every time the manufacturing process changes the simulation is to be performed again to achieve a proper understanding of the statistical variations.

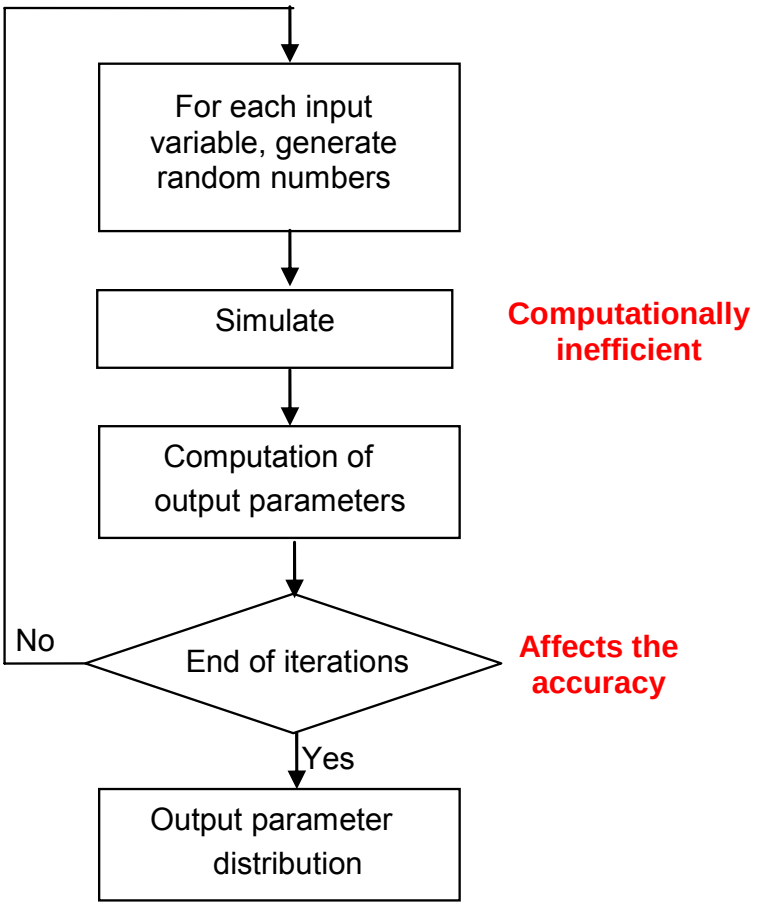


Fig. 5.1 Monte Carlo Analysis simulation flow.

Another approach to analysing the effect of process variation is to use Design of Experiment (DOE) and Response Surface Modelling (RSM) techniques to

determine process combinations that should be simulated to achieve the corresponding response. The resulting response surfaces are then used to statistically produce process distributions that result from control factor variations [13-20]. A comparison between the efficiency of Monte Carlo Analysis and the DOE technique in analysing the effect of parameter variations on the threshold voltage of n- and p-MOS has been discussed in the literature [21]. It has been observed from [21] that the threshold voltage distributions obtained using Monte Carlo technique and DOE/ RSM approach analyses fits well with simulations lasting for a few seconds in the case of DOE/ RSM approach compared to 10 minutes for Monte Carlo technique. The analyses in this Chapter is limited to DOE/RSM approach due to its computational efficiency compared to Monte Carlo technique and the large number of process parameters in the technology to be analysed [12, 13].

CAD tools are essential to perform optimisation and the creation of the models to accurately predict the effects of process variability, thereby increasing the overall productivity of the designs. These CAD tools were discussed in Chapters 2 and 3. The entire work related to the analysis of variability is performed by developing a standard Si and strained-Si technology library by simulations using TSUPREM4 [8] (process simulator), MEDICI [9] (device simulator), AURORA [22] (parameter extraction program), MINITAB (Design of experiments) [23] and PSPICE [10] (circuit simulator) subjected to different operating conditions and varying amounts of strain in the channel. High performance (HP) strained-Si and standard Si devices on a 65nm technology node are considered for the analyses. ITRS/MASTAR [1, 24] device parameters have been used for the device and circuit dimensions and operating conditions for simulation.

The ensuing discussion will outline the methodology used to identify the process parameters of a typical 65nm technology node and the methodology to study the impact of these process parameters on circuit output response. The demonstration circuits used for the analyses will comprise two basic asynchronous logic building blocks, namely, the Mutual Exclusion Element (MUTEX) and Jamb Latch synchronizer. In the analysis the significant parameters which impact on the circuit outputs are identified and the response surface modelled. Finally a comparison is made between strained and conventional Si implementations of the

demonstration circuits with respect to an important design parameter, namely, "metastability resolution time".

Apart from the advantages of using strained-Si technology in high-speed low-power digital and analog applications which were discussed in Chapter 4, the results obtained from the analyses of MUTEX and Jamb Latch based on strained-Si technology indicates that strained-Si technology combined with the advantages of asynchronous design techniques is a viable contender as an implementation technology for future circuit applications.

5.2 Identification of Process Parameters

The Monte Carlo analysis was the generally adopted approach used to study the effects of parameter variability in a wide range of applications for example, chemical industries, flight simulations etc; where the number of parameters are limited [5-7]. However, when the number of parameters considered increases, this approach, although accurate, is computationally inefficient. The technique described in this chapter to identify semiconductor process parameters whose variability would impact most on the circuit characteristics, considered critical by the designer, is realised through a two phase process using DOE and RSM statistical tools [13]. DOE is a strategy used to create a set of experiments in which the range of variables can be altered systematically to enable any correlation between variables to be ascertained to determine the main contributing factors to the variation in, for example, design parameters of interest. In this particular application the 'experiments' in DOE are simulations. Within the DOE procedure the Plackett-Burman screening technique [13, 23] is incorporated to screen the most significant parameters which will subsequently be used in the second phase to model the response surface.

Experimental Design formally represents a sequence of experiments to be performed, expressed in terms of factors or design variables set at specified levels that is, predefined values. Having identified the most influential parameters it is essential to reflect the effects of the variations onto the output response of the

system under investigation. In this instance the 'Variations' are the semiconductor process parameters and its effect on some aspect of the output response of a circuit which is of interest.

Variability analysis is performed by developing a standard Si and strained-Si technology library. In creating a technology library a number of CAD tools are used in sequence as shown in Figure 5.2; namely TSUPREM4, MEDICI, AURORA and PSPICE, DOE analysis is performed using the statistical package MINITAB [23].

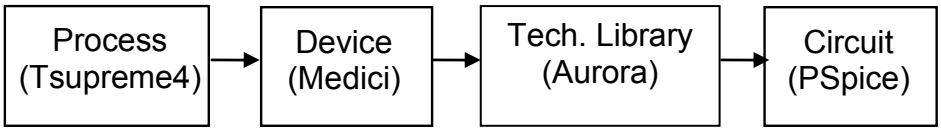


Fig. 5.2 CAD simulator flow.

5.3 Methodology to Study the Impact of Process Variations on Circuit Parameters

As a simple demonstration of the above procedures, a Mutual Exclusion Element (MUTEX) [25] and a Jamb Latch synchronizer [26, 27] were selected to study the impact of process variations on particular circuit parameters. The MUTEX and Jamb Latch circuits are important building blocks which are used extensively in asynchronous circuits [28]. In the design of a MUTEX two critical timing parameters, t_{off} and t_m, which are discussed in detail in Section 5.4 are important for the correct operation of the circuit, they are also susceptible to the effects of process variations; whereas in the case of Jamb Latch, t_m is the output response on which the process variations are affected.

The implementation technology chosen for the demonstration circuits is strained-Si. The amount of strain in the strained-Si device can be considered as a process parameter and hence gives an easy way of comparing the performance of the circuits implemented in the standard silicon process technology (0% strain) and one implemented in the strained-Si technology.

The electrical characteristics of strained-Si and conventional Si n- and p-MOSFETs are applied to AURORA parameter extraction simulator for optimization

and extraction of PSPICE models (the parameter extraction strategy and the models are shown in Appendix B). The strained-Si and Si PSPICE transistor library based on different strain conditions and technology nodes were developed for the analyses. Subsequently these transistors were used in the design of the demonstration circuits in PSPICE and simulations were performed to evaluate the variations of t_m and t_{off} of strained-Si and Si based MUTEX circuits under different strain and process conditions.

The procedure followed to study the impact of process variations on the chosen circuits is shown in Figure 5.3 in which there are two paths. The result from the first path is the benchmark for the circuit, and the second is the stochastic output based on the variability of the process and device parameters using DOE. The best process parameters can then be selected, depending upon the application of the circuit, to ensure a good yield.

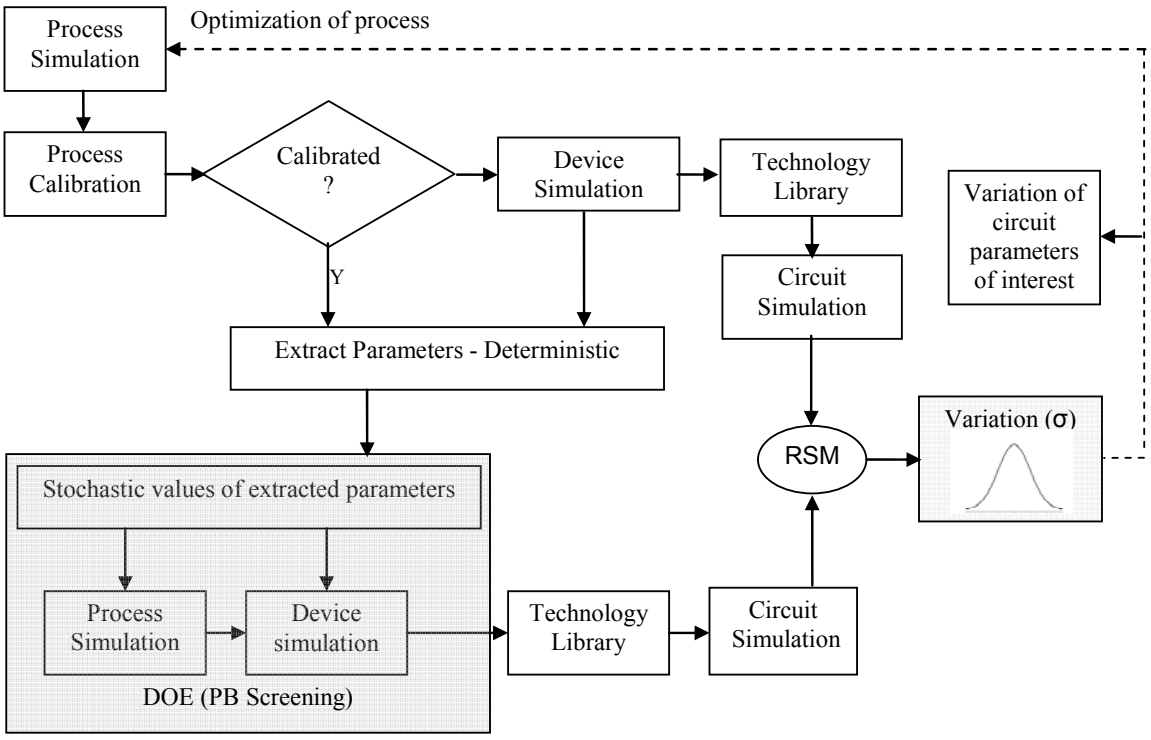


Fig. 5.3 Methodology to study the effect of process variations.

To study the impact of the variability of process parameters on circuit characteristics the main process parameters were selected from a chosen 65nm

process (with ITRS specifications). State of the art 65nm technology was selected as the technology node of choice as its variability is seen to be larger in comparison to 300nm technology node [29]. Monte Carlo Analysis would have been the obvious choice to study the effects of variability as it is the conventional statistical approach when considering large amounts of data. However, to study the effects of the variation of large number of processing parameters on a circuit a huge number of process, device and circuit simulations would need to be performed which is highly complicated, computationally inefficient and time consuming but accurate using the Monte Carlo approach. This is due to the fact that the accuracy of the results from Monte Carlo Analysis depends on the number of random samples, a large number of finite samples are necessary to achieve reasonable accuracy for a complex system which was shown in Figure 5.1. Consequently, a different statistical method called DOE was used which when compared to Monte Carlo approach is very efficient computationally [13]. Figure 5.4 shows the steps undertaken in the analysis of the effects of process variation using DOE/RSM.

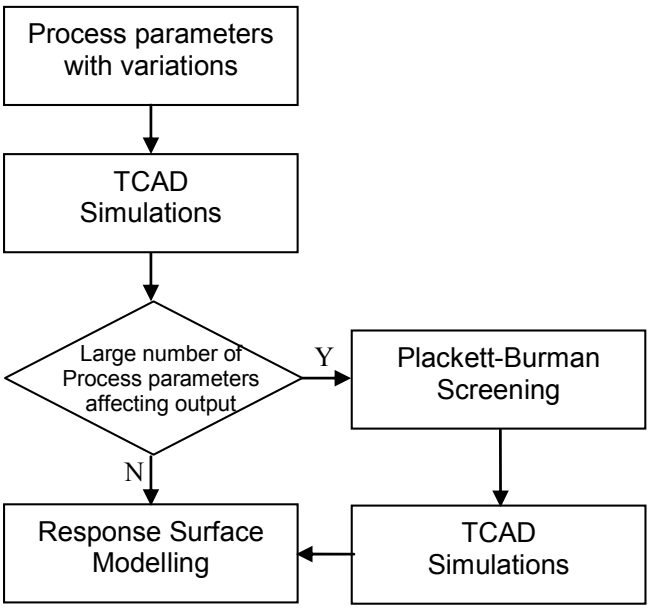


Fig. 5.4 Flow chart of the variability analysis.

The DOE method involves screening out insignificant process parameters and modelling the significant parameters for variability analysis. Out of seventeen

parameters identified from a typical 65nm process technology [29], eight parameters are shown in Table 5.2 which seem to be important and uncontrollable at different process steps.

For the variability analysis, the process parameters were all varied over a range of 10% from their nominal values. However, the temperature was varied as $\pm 5^{o}C$ from the nominal. If Monte Carlo analysis was used in this instance, approximately 1000 simulations would have been required to achieve a reasonable accuracy to model the parameters assuming that around 100 simulations are required for each parameter (eight parameters short listed in Table 5.2). When two phase DOE/RSM method is applied, with the help of screening, the number of simulations can be dramatically reduced as the number of simulations for modelling using DOE could not exceed $2^{n}+2n+1$, where n is the number of significant parameters [13, 23, 29]. Parameter screening is achieved using Plackett-Burman (PB) method [29].

Symbol	Description	Unit	Mean Value
x_1	Gate length	nm	25
x_2	Substrate doping	Cm^{-3}	8.5e18
x_3	S/D doping	Cm^{-3}	2.7e20
x_4	LDD doping	Cm^{-3}	1e20
X_5	Oxide thickness	nm	0.9
X_6	Junction depth	nm	28
X_7	LDD junction depth	nm	14
X_8	RTA temp.	^{o}C	1000

Table. 5.2 Parameters identified in a typical 65nm technology process [30].

Among the list of parameters shown in Table 5.2, the effective gate length (L_{eff}) is not considered separately, as L_{eff} depends on substrate doping concentration and other process conditions. The significant parameters are chosen based on a

threshold value, decided by the user, below which it is considered that the parameter is insignificant. The amount of strain was also considered as an important parameter, hence it was added to the significant parameters obtained after screening to model the variability using RSM; this permitted a comparison between silicon and strained-Si technologies to be made.

5.4 Analysis of circuit design parameters with process variability

This section discusses the impact of process parameter fluctuations on the output variables of interest for the MUTEX and Jamb Latch synchronizer realised with strained-Si technology. These circuits are also realised with standard Si technology to see the advantages of using strained-Si technology in asynchronous circuits.

5.4.1 Mutual Exclusion Element (MUTEX)

Asynchronous communication systems require its input channels to be mutually exclusive. In this type of systems there are several situations where a resource is shared between several independent processes. Instrumental in this sharing process is a basic circuit called a MUTEX whose block diagram is shown in Figure 5.5.

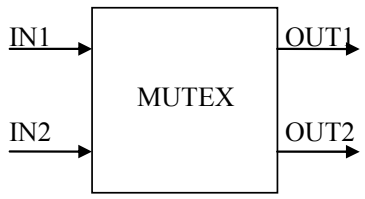


Fig. 5.5 Block diagram of a MUTEX with two input signals.

The input signals IN1 and IN2 shown in Figure 5.5 are two signals that originate from two independent sources, and the task of the MUTEX is to pass these inputs to the corresponding outputs OUT1 and OUT2 in such a way that only one of the outputs is active at any given time [25]. However in the case where there is only one input request which is active then the operation is insignificant. If the first input request arrives earlier than the second one, the latter request is blocked until the first request is de-asserted. However there is a problem when both the input signals IN1

and IN2 are asserted at the same time. In this instance the MUTEX has to decide upon which input should be given priority; the length of the time taken to resolve the situation is called 'metastability resolution time' denoted by t_m. A similar situation can arise in synchronous memory element when data and clock signals violate setup and hold times. The minimum separation time between the input signals to ensure stable operation of the MUTEX is called the offset time t_{off}. Both t_m and t_{off} are critical parameters in the design of a MUTEX and are affected by process variations.

The implementation of the MUTEX is shown in Figure 5.6 involves a pair of cross coupled NAND gates and a metastability filter.

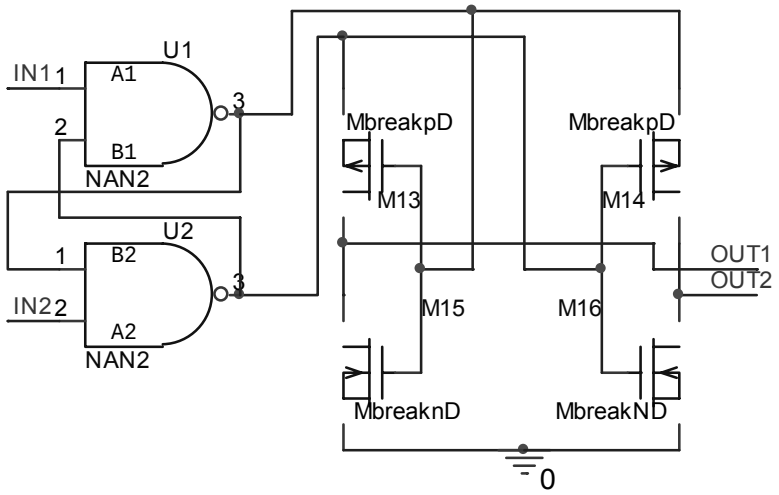


Fig. 5.6 MUTEX implemented using cross coupled NAND gates with metastability filter.

The cross coupled NAND gates enable one input to block the other. If both inputs IN1 and IN2 are asserted at the same time, the circuit becomes metastable with both signals U1 and U2 sitting at a voltage halfway between the supply voltage and ground. The metastability filter prevents these undefined values from propagating to the outputs; OUT1 and OUT2 are both kept low until signals U1 and U2 differ by more than a transistor threshold voltage. The metastability filter can be implemented using two buffers whose logic thresholds have been made particularly high (or low) by trimming the strengths of the pull-up and pull-down transistor paths.

5.4.1.1 Variability analysis: MUTEX

The flow chart of the analysis to study the variability of the process parameters in a MUTEX realised with strained-Si and standard Si technology is shown in Figure 5.7. The most significant parameters have to be identified using Plackett-Burman screening method. The significant parameters are chosen based on a threshold value below which it is considered that the input parameter is insignificant. Considering strain as an important parameter, it is added with the significant parameters obtained after screening to model the variability using RSM. Here the widths of transistors in MUTEX were kept constant to study the variation of t_m and t_{off} based on process variation. Even though the variation of widths of transistors in MUTEX is very important as it is typical for different applications using MUTEX, it is not considered here for the analyses.

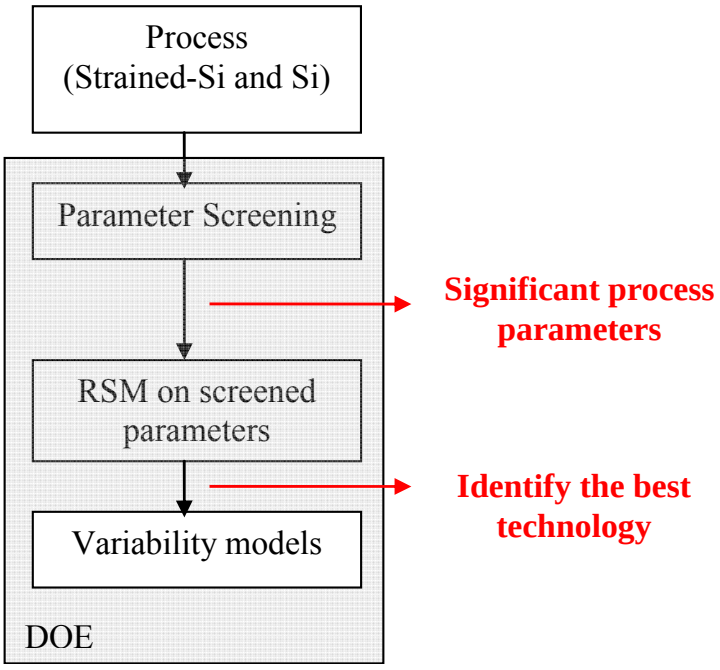


Fig. 5.7 Flow chart of the variability analysis of MUTEX.

5.4.1.2 Impact of Process variation on the behaviour of the MUTEX

The transistors used to implement a MUTEX are subjected to process variability conditions to study its effects on t_m and t_{off} using DOE/RSM methods.

The supply voltage is fixed at 1V. The first phase of the variability analysis is to find out the significant process parameter (and its variation) which has the greatest impact on t_m and t_{off}. All the process parameters shown in Table 5.2 were varied by ± 10%. However, the amount of strain is varied from 0% (conventional Si technology) to 0.99% (strained-Si). From the parameter screening technique (Placket-Burman method), Oxide thickness (t_{ox}), N_{sub} (substrate doping) and the amount of strain are found to be the most significant parameters which impact mostly on t_m and t_{off} [30]. RSM is then performed for the three parameters. Since only three parameters are identified as being the most significant parameters in the case of MUTEX after the screening technique, only fifteen (2^n+2n+1) simulations were required to generate the response surface model of the MUTEX outputs [13, 23]. The statistical p-values were obtained from the analysis. P-values provide the way of testing the relationship between the predictor (input variables which are significant process parameters) and response (output variables which are t_m and t_{off}) [13, 21] and are used to determine statistically significant terms in the model. The p-values for linear, square and interaction effects obtained by surface modelling are shown in Table 5.3.

Effects	**Process parameter**	**t_m**	**t_{off}**
		p- values	**p- values**
Linear	t_{ox}	0.454	0.746
”	N_{sub}	0.140	0.016
”	strain (%)	0.008	0.196
Square	t_{ox} x t_{ox}	0.989	0.742
”	N_{sub} x N_{sub}	0.999	0.800
”	Strain (%) x strain (%)	0.028	0.181
Interaction	t_{ox} x N_{sub}	0.580	0.336
”	t_{ox} x strain (%)	0.912	0.628
”	N_{sub} x strain (%)	0.004	0.010

Table. 5.3 The linear, square and interaction effects of process parameters on t_m and t_{off} of MUTEX.

The terms with p-values ≤ 0.5 are statistically significant [13]. From the linear effects it can be seen that strain is the most significant process parameter which impacts t_m, and for t_{off} it is $N_{sub.}$ For square effects; the strain value has the highest impact on t_m and t_{off} of the MUTEX. However, among the interaction effects the interaction of N_{sub} and strain are also significant parameters for t_m and t_{off} variation.

Figure 5.8 shows the surface plot of the impact of interaction effects of N_{sub} and strain on t_{off} of a MUTEX. It can be seen that t_{off} is giving a large range of variation for 0% strain (conventional Si) when N_{sub} is changing from -10% to +10% which is undesirable. When the strain in the channel of the transistors is increasing it can seen that the variation of t_{off} with N_{sub} is not high compared to the lowest strain (0%) which gives an edge for the strained-Si technology over the conventional Si. This can be attributed to the enhancement of the mobility of electrons and holes are not the same in the case of strained-Si [30-32] and hence the transconductance ratio and the change in threshold voltage is not the same as in conventional Si. As mentioned above the strain is found to be the most significant parameter which impacts on t_{off}. This is advantageous due to the fact that the amount of stress applied to the channel to create strain can be controlled by the Ge% in the wafer (in the case of bi-axial strain) or by the process steps in local strain (uni-axial strain).

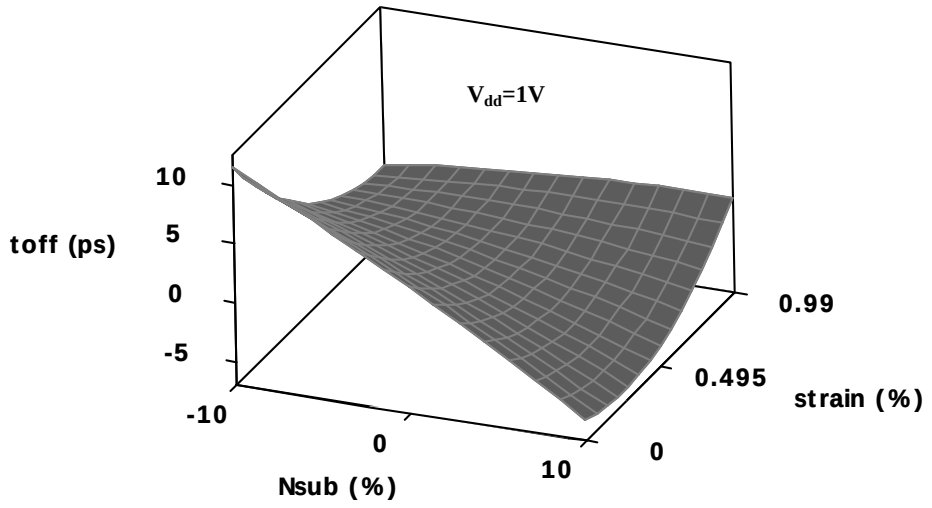


Fig. 5.8 Surface plot of the impact of t_{off} due to the variation of N_{sub} by 10% and strain varying from 0.0% to 0.99% of MUTEX.

The surface plot to show the impact of interaction effects of N_{sub} and strain on t_m of MUTEX is shown in Figure 5.9. Similar to t_{off}, t_m is also giving a large range of variation for 0% strain (conventional Si) when N_{sub} is changing from -10% to +10%. t_m of MUTEX is found to decrease when the amount of strain applied in the channel of transistors is increasing. It can also be seen that t_m is almost 100% less when N_{sub} is +10% from the nominal at 0.99% strain compared to the lowest strain (0%) which again indicates that strained-Si technology is a better candidate to realise a MUTEX compared to the conventional Si MUTEX.

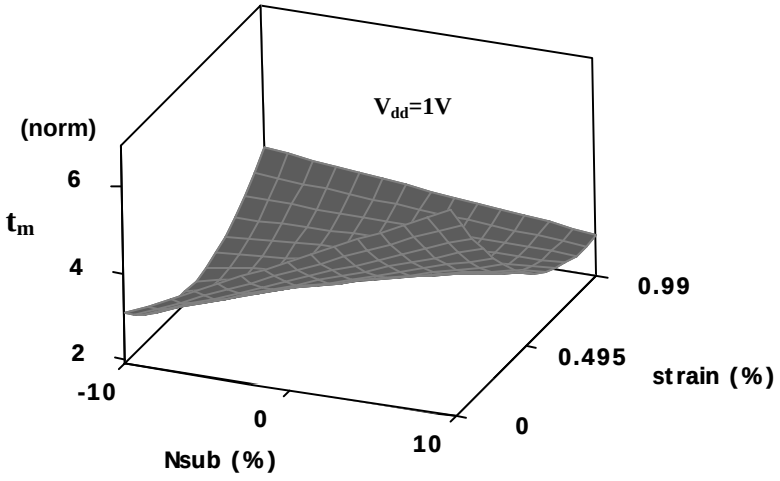


Fig. 5.9 Surface plot to show the impact of t_m due to the variation of N_{sub} by 10% and strain varying from 0.0% to 0.99% of MUTEX.

Table 5.4 shows the validity of the model developed to study the impact of process variation on t_{off} and t_m using DOE/ RSM methods. The validity of the model can be checked using the R-square statistics. R-square is known as the coefficient of determination and is a statistical measure of how well the regression line approximates to the real data points. R-square and adjusted R-square statistics represent the amount of variation in the response that is explained by the model. The R-square value always increases with the addition of a variable to the model, regardless of whether the additional variable is statistically significant or not. Hence, models with large R-square values give poor predictions of new observations or estimates of the response. Adjusted R-square is a modified R-square for the number of terms in the model. Unlike R-square, adjusted R-square may get smaller when

unnecessary or additional terms are added to the model. It can be seen from Table 5.4 that the adjusted R-square value is less for both t_m and t_{off}.

	R-sq	R-sq (adj)
t_m	93.36%	81.41%
t_{off}	87.49	64.98%

Table. 5.4 The validity of the model.

5.4.2 Jamb Latch Synchronizer

Whenever there is a signal transfer between two systems operating at different frequencies or the same frequency with different phases, the signals coming from the first system become asynchronous to the second system thereby increasing the probability of the output of first system going into metastability and propagating the asynchronous signal into the second system [33]. The most common approach to minimize the metastability problem is to use a synchronizing circuit to take the asynchronous signal and align it to the timing regimen of the rest of the system. The Jamb Latch is a simple circuit commonly used as a synchronizer because of its relatively better performance [28] compared to the conventional two stage synchronizers; its basic configuration is shown in Figure 5.10.

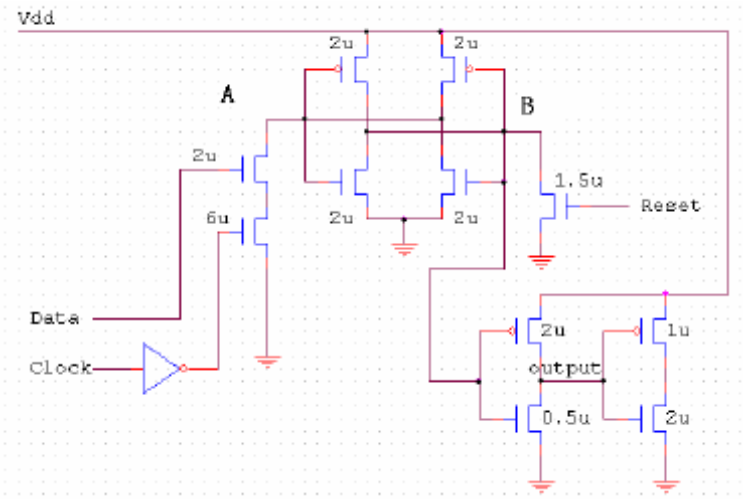


Fig. 5.10 Schematic of Jamb Latch [26].

5.4.2.1 Variability Analysis: Jamb Latch Synchronizer

In this section we explore the possibility of performance enhancement in terms of metastability resolution time in synchronizers using strained-Si transistors. The analysis is based on the Jamb Latch which is a commonly used synchronizer in real time circuits [26, 27]. Although the studies are based on the behaviour of the Jamb Latch synchronizer using strained-Si as the base material, emphasis is given to the comparison of strained-Si technology against that of conventional Si technology with similar device dimensions. The device dimensions and other physical parameters for DSM high performance (HP) devices are taken from *MASTAR* [24].

The electrical characteristics of strained-Si and conventional Si n- and p-MOSFETs are applied to AURORA parameter extraction simulator for optimization and extraction of PSPICE models. These transistors are configured for the design of Jamb Latch in PSPICE. Simulations are performed to evaluate the metastability resolution time of strained-Si and Si based Jamb Latch under different operating conditions and dimensions. To study the impact of variability of the process parameters on t_m, DOE/RSM is applied to develop corresponding variability models [13]. Plackett- Burman screening is performed to identify the most significant process parameters which have the greatest influence on t_m. The amount of strain in the channel, the operating temperature and the supply voltage are found to be the most significant parameters affecting the metastability resolution time of the Jamb Latch.

The ratio of widths of p- MOS to n- MOS in Jamb Latch is taken to be 1μm:1μm, which is considered as the best as reported [26]. The simulations are performed at supply voltages ranging from 0.9V to 1.3V for the transistors under different process conditions. The supply voltage is limited to the above range as is specified by ITRS for the HP devices in 65nm technology node.

The variation of t_m with respect to the various technology nodes is plotted in Figure 5.11 for supply voltages of 1.1V and 1.2V. It can be seen from Figure 5.11 that t_m reduces with scaling of devices.

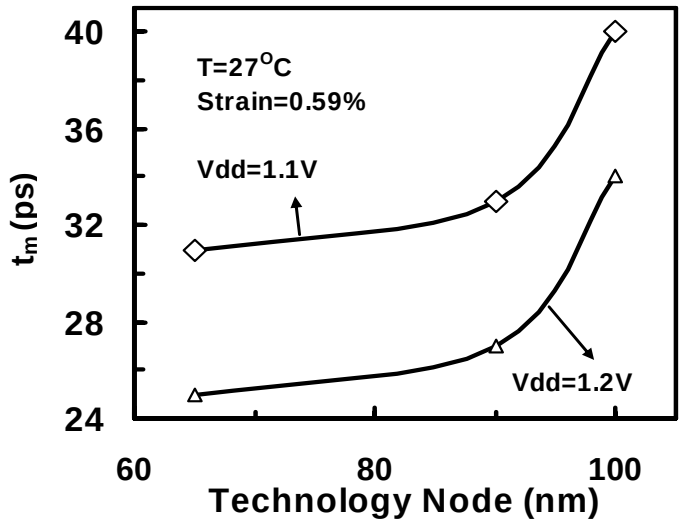


Fig. 5.11 Metastability resolution time Vs technology node for strained-Si CMOS based Jamb Latch at 1.1V and 1.2V.

Figure 5.12 shows the variation of t_m against the amount of strain applied in the channel of transistors in same technology node (65nm). For comparison, supply voltages of 1.1V and 1.2V are considered. It can be seen from the figure that the resolution time decreases almost linearly with the increase in the amount of strain applied to the transistors that is, standard Si devices (with 0% strain) are taking more time to resolve the metastability problem compared to the strained-Si devices. The value of t_m is almost 66% less for strained-Si based Jamb Latch compared to that of standard Si Jamb Latch. These results are attributed to the fundamental differences between standard Si and strained-Si technology. The strained-Si device inherently has higher transconductance and lower threshold voltage compared to that of the standard Si due to the presence of strain and smaller band gap. However, it is not expected that the same trend to continue with the increase in the amount of strain.

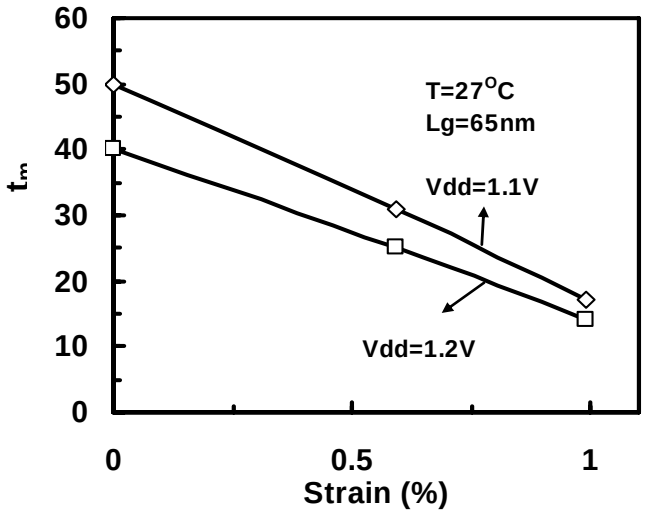


Fig. 5.12 Metastability resolution time Vs amount of strain in the channel for supply voltages of 1.1V and 1.2V at 65nm technology node.

Figure 5.13 shows the comparison of t_m of the Jamb Latch realized with transistors with 0.99% and 0.59% amounts of strain in the channel.

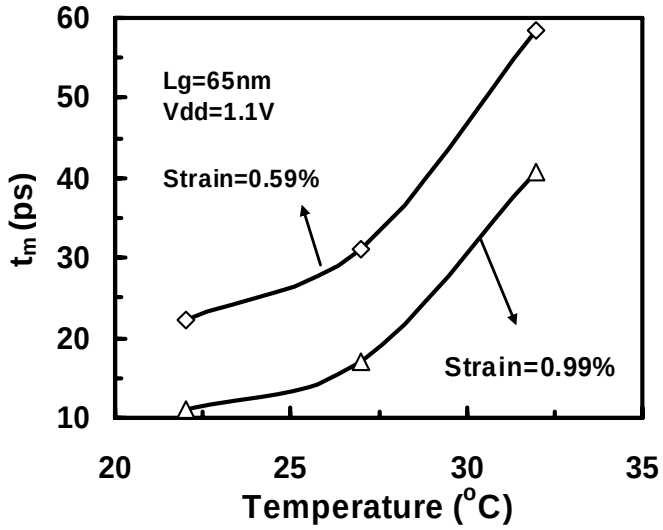


Fig. 5.13 Variation of t_m against temperature for transistors with different amount of strain in the channel.

A supply voltage of 1.1V is given for the circuit operation. It can be seen from the figure that the strained-Si with 0.99% strain in the channel exhibits a lower metastability time compared to that of strained-Si Jamb Latch with 0.59% for different ranges of temperature. It can also be seen from figure that t_m reduces with the reduction in temperature. This again indicates that strained-Si is a better candidate compared to Si technology for asynchronous circuit applications.

RSM has been used to study the impact of the variation of the amount of strain, supply voltage and temperature on t_m. Figure 5.14 shows the response surface plot of t_m (ps) as the function of nominal temperature and power supply voltage. The impact of temperature is evident from the response surface. At the low levels of temperatures and supply voltages, t_m has decreased and has its maximum value at the high level of temperature and lower level of supply voltage. Hence the response surface plot shows the range of variations caused in t_m due to the variations in the electrical parameters.

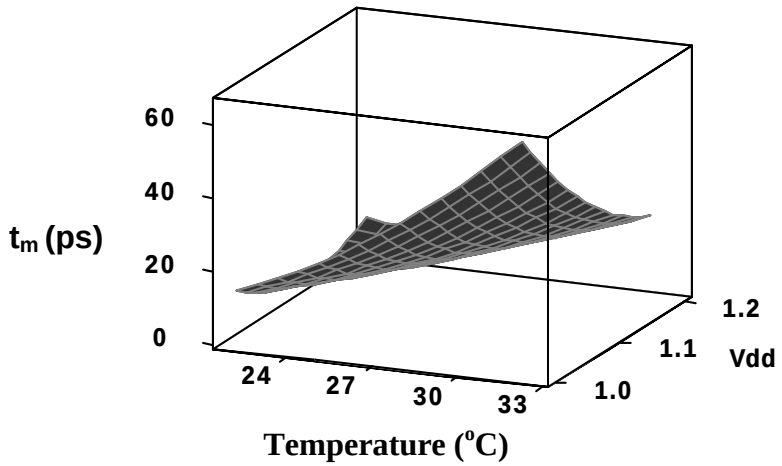


Fig. 5.14 Response surface plot of t_m as function of operating temperature and supply voltage.

Figure 5.15 shows the response surface of t_m with respect to amount of strain and supply voltage. The total change in t_m is about 40ps due to the variations of the strain applied in the channel and the supply voltage. Metastability resolution time is lower with greater amounts of strain and lower values of supply voltage. It can be

seen from the figure that the impact of the amount of strain is more significant on t_m than the supply voltage variation.

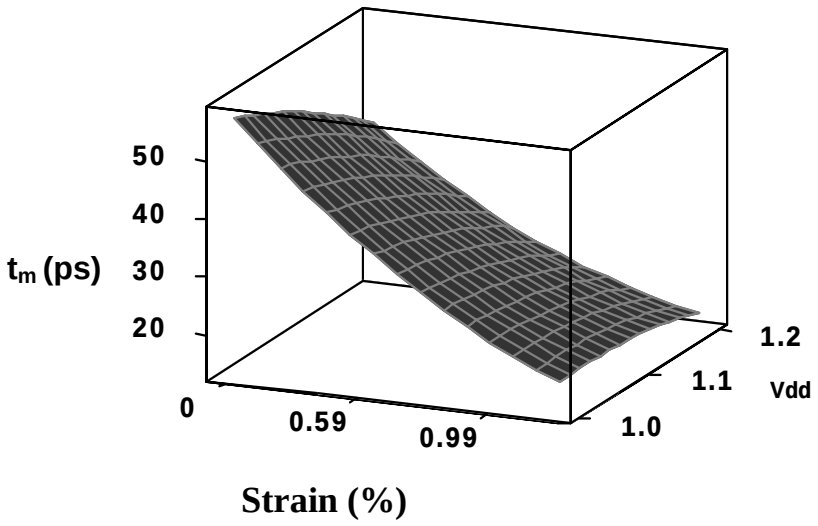


Fig. 5.15 Response surface plot of t_m as function of amount of strain and supply voltage.

The results obtained from the analyses of MUTEX and Jamb Latch indicate the potential and flexibility of strained-Si technology for asynchronous circuit applications. Strained-Si is found to be a better technology than standard Si for making synchronizers. Besides, it is possible to achieve significantly better low-power, high-speed, overall noise margin and better timing closure performance from strained-Si based circuits than conventional Si circuits. These facts coupled with the availability of higher supply voltage reduction window makes strained-Si based circuits an ideal candidate for futuristic synchronous and asynchronous circuit design. Variability analysis using DOE to study the variation of metastability resolution time with process and operational parameters (V_{dd} and nominal temperature) in three dimensional spaces indicates that the variation of strain in the channel is found to be a significant parameter compared to the supply voltage variation. This information can be used by the designer to accommodate the variations hence ensuring real time operations without fault.

5.5 Summary

In this Chapter a methodology has been explored to study the impact of process variability on circuits. This Chapter also analyses the potential and flexibility of strained-Si technology for asynchronous circuit applications. The method is applied to Jamb Latch synchronizer and MUTEX which are the basic building blocks of many modern asynchronous circuits. The state of the art 65nm high performance transistors are used for the analyses. A screening technique is used to screen the insignificant parameters from the set of physical and dimensional parameters involved in the typical 65nm technology node. For the 65nm technology node, simulations have been performed at process, device and circuit level to study the impact of process variations. However, it would have been computationally inefficient in the case of Monte Carlo technique which is the conventional method to study the variations. Interaction of process parameters in three dimensional spaces has been studied. The variation of strain in the channel is found to the most significant process parameter in case of strained-Si technology.

Overall:-

- MUTEX and Jamb Latch synchronizer using transistors build using strained-Si and conventional Si technologies are analysed and compared.
- By noting the metastability resolution time and offset time of MUTEX, strained-Si technology is seen to be a better alternative to realising asynchronous circuits than Si technology. It has also been observed that the t_m is almost 100% less when N_{sub} is +10% from the nominal at 0.99% strain compared to the lowest strain (0%) which again indicates that strained-Si technology is a better candidate for building MUTEX compared to the conventional Si technology.
- The metastability resolution time of Jamb Latch synchroniser using transistors realized with strained-Si technology is seen to be less than that with conventional Si technology. This in effect increases the Mean Time Between Failure (MTBF) [33]. The t_m is around 66% less for strained-Si based Jamb Latch compared to that of standard Si Jamb Latch.
- Variability in IC manufacturing and its impact on circuits are reviewed.

- Advantages of using statistical techniques to understand variability rather than corner case methods are discussed.
- The algorithm to model the variability has been extensively discussed.
- DOE/ RSM have been introduced to analyse variability in MUTEX and Jamb Latch synchroniser circuits.

References

[1] International Technology Roadmap for Semiconductors, "Process Integration, Devices and Structures", 2007 Ed.

[2] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-Power CMOS Digital Design", *IEEE Journal of Solid-State Circuits*, vol. 27, pp. 473-484, 1992.

[3] L. Nazhandali, B. Zhai, A. Olson, A. A.-R. Reeves, A., et al., "Energy optimization of subthreshold-voltage sensor network processors", *International Symposium on Computer Architecture*, 2005.

[4] B. P. Wong, *Nano-CMOS Circuit and Physical Design.* New Jersey: John Wiley, 2005.

[5] S. R. Nassif, A. J. Strojwas, and S. W. Director, "FABRICS II: A Statistically Based IC Fabrication Process Simulator", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 3, pp. 40-46, 1984.

[6] K. K. Low, "A methodology for statistical integrated circuit process design", Thesis (Ph.D), Carnegie Mellon University, 1989.

[7] R. Cartuyvels, R. Booth. L. Dupas, and K. De Meyer, "Process technology optimisation using an integrated process and device sequencing system", *ESSDERC'92*, pp 507-510, 1992.

[8] TSUPREM4 user manual: *Two-Dimensional Process simulation program.* Sunnyvale: Synopsys Inc., 2000.

[9] MEDICI user manual: *Two-Dimensional Device simulation program.* Sunnyvale: Synopsys, Inc., 2000.

[10] PSPICE user manual: *ORCAD PCB design suite.* San Jose: Cadence design systems Inc., 2006

[11] S. Tirumala, Y. Mahotin, X. Lin, V. Moroz, et.al., "Bringing Manufacturing into Design via Process-Dependent SPICE Models", *International Symposium on Quality Electronic Design (ISQED '06)*, 2006.

[12] C. Jacoboni, and P. Lugli, *The Monte Carlo method for semiconductor device simulation.* New York: Springer, 1989.

[13] D. C. Montgomery, *Design and Analysis of Experiments.* New York: Wiley, 1991.

[14] M. Cecchetti, M. Lissoni, C. Lombardi, and A. Marmiroli, "Process analysis using RSM and simulation", *ESSDERC'92*, pp. 511-514, 1992.

[15] R. Cartuyvels, R Booth, S. Kubicek, L. Dupas, and K. De Meyer, "A Powerful TCAD System Including Advand RSM Techniques for Variaus Engineering Optimization Problems", *SISDEP'93*, pp.29-32, 1993.

[16] A. J. Walton, U. Fallon, M. L. Newsam, R. S. Ferguson, and D. Sprevak, "Using RSM Techniques to Contour Plot Response Distributions of Semiconductor F'rocesses", *ESSDERC'94*, 1994.

[17] G. J. Gaston, and A. J. Walton, "The integation of simulation and response surface methodology for the optimisation of IC processes", *IEEE Transactions on Semiconductor Manufacturing*, Vol. 17, pp 511-514, 1994.

[18] A. R. Alvarez, B. L. Abdi, D. L. Young, H. D. Weed, et.al., "Application of statistical design and response surface methods to computer-aided VLSI device design", *IEEE Trans. Computer-Aided Design*, vol. 7, pp. 272–287, 1988.

[19] S. Nassif, "Design for manufacturability in DSM technologies", IEEE *International Symposium on Quality Electronics Design*, pp. 451-454, 2000.

[20] H. C. Srinivasaiah, and N. Bhat, "Implant Dose Sensitivity of 0.1μm CMOS Inverter Delay", *International conference on VLSI Design (VLSID'02)*, 2002.

[21] X. Gan and A. J. Walton, "A comparison of direct Monte-Carlo and DOE simulations for optimising IC processes", *International Conference on Solid-State and Integrated Circuit Technology*, 1995.

[22] AURORA user guide: *AURORA simulation program.* Sunnyvale: Synopsys Inc., 2006.

[23] Minitab, *Minitab Release 15, Statistical Software*. Minitab Inc., 2006.

[24] Model for assessment of cmos technologies and roadmaps, *MASTAR.* ST Microelectronics, Advanced devices research group, 2000-2005.

[25] J. SparsÃ¸ and S. B. Furber, *Principles of Asynchronous Circuit Design: A Systems Perspective*. Boston: Kluwer Academic Publishers, 2001.

[26] J. Zhou, D. Kinniment, G. Russell, and A. Yakovlev, "A robust synchronizer", *IEEE Computer Society Annual Symposium on Emerging VLSI Technologies and Architectures*, 2006.

[27] C. Dike and E. Burton, "Miller and noise effects in a synchronizing flip-flop", *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 849-855, 1999.

[28] A. M. Abas, A. Bystrov, D. J. Kinniment, O.V. Maevsky et.al., "Time Difference Amplifier", *IEEE Electronic Letters*, Vol. 38, pp. 1437-1438, Nov. 2002.

[29] S. Shedabale, S. Chattopadhyay, H. Ramakrishnan, S Uppal, et.al., "Statistical Analysis and Variability Modelling for 0.3um n-MOS Si Process", *Post Graduate Conference*, Newcastle University, 2005.

[30] M. T. Currie, C. W. Leitz, T. A. Langdo, G. Taraschi, and E. A. Fitzgerald, "Carrier mobilities and process stability of strained-Si n and p-MOSFETs on SiGe virtual substrates", *Journal of Vacuum Science and Technology*, Vol. 19, pp. 2268–2279, Nov. 2001.

[31] S. Chattopadhyay, L. D. Driscoll, K. S. K. Kwa, S. H. Olsen, et.al., "Strained Si MOSFETs on relaxed SiGe platforms: performance and challenges", *Solid-State Electronics*, vol. 48, pp. 1407-1416, 2004.

[32] K. Rim, J. L. Hoyt, and J. F. Gibbons, "Fabrication and analysis of deep submicron strained-Si n-MOSFET's", *IEEE Transactions on Electron Devices*, vol. 47, pp. 1406-1415, 2000.

[33] J. F. Wakerly, *Digital design : Principles and Practices*. New Jersey: Prentice Hall, 1990.

CHAPTER 6

DISCUSSION, CONCLUSIONS AND FUTURE WORK

6.1 Discussion

The work in this thesis explores the potential of using strained-Si as an implementation technology for higher performance low-power circuit applications such as the portable consumer market, which is one of the principal economic drivers for the semiconductor industry [1]. The investigation is undertaken from two view points namely:

- Technology Perspective: Focuses on the improved carrier mobility in strained-Si devices and the impact this has on the device/circuit performance compared to conventional Si devices.
- Design Perspective: Investigates how the methods of reducing power consumption in circuits such as supply voltage reduction and device stacking effect the performance of strained and conventional Si devices.

Furthermore, as device dimensions are reduced the effect of process variations on device performance becomes problematic, consequently the susceptibility of strained-Si circuits to process variation is analysed using the statistical methods called Design Of Experiments (DOE) and Response Surface Modelling (RSM).

When the simulators in the TCAD suite were used similar processing and operating conditions were assumed for both strained and conventional Si technologies, this enabled comparisons to be easily made. The parameters used for device simulation were presented in Chapter 3. As summarized in Chapter 3 the advantages of using strained-Si technology over conventional Si technology to build the transistor are:

- The saturated drain current for the strained-Si n-MOSFET (0.99% strain in the channel) indicates a 21% increase over that of the conventional Si counterpart.
- An increase of 28% in drain current is observed in the case of strained-Si p-MOSFET (0.99% strain in the channel) compared to the Si p-MOSFET.

Furthermore the disadvantages associated with bi-axial strained-Si technology for example, higher subthreshold leakage current and self-heating which is due to the lower thermal conductivity of underlying $Si_{(1-x)}Ge_x$ were analysed. The parameters affected by the phenomenon of self-heating for example, mobility, subthreshold leakage current and threshold voltage were also discussed in Chapter 3.

The disadvantages of transistors built using bi-axial strained-Si technology were observed in Chapter 3 are summarized:

- The subthreshold leakage current of strained-Si based transistor was seen to be higher by one order of magnitude compared to the standard Si counterpart.
- Due to self-heating the saturation current at higher drain and source voltages was reduced by around 4%.

- Self-heating affects the threshold voltage of transistor. This may dramatically change the timing behaviour of circuits when implemented using strained-Si technology which in turn changes the performance of a digital circuit. However the results mentioned in Chapter 3 indicate that the effect of self-heating can be reduced by using a thin virtual substrate.

The advantages discussed above are exploited in making strained-Si circuits to be used for high-speed low-power digital circuit applications by using some techniques to compromise the disadvantages.

As outlined in Chapter 4 the advantages obtained from a strained-Si implementation of the inverter circuit are:

- Strained-Si inverter was shown to be faster than the standard Si inverter by 26%. The speed margin will increase with deeply cascaded circuits.
- The advantage of strained-Si in terms of speed was also demonstrated by creating a ring oscillator (series of cascaded inverters) which was 17% faster than its standard Si counterpart.
- When the supply voltage of strained-Si inverter is reduced the same speed can be achieved as that of Si inverter but with very low power dissipation or alternatively with similar power dissipation to Si inverter, a higher performance can be achieved; this is due to strained-Si having a lower threshold voltage and higher mobility than conventional Si.

The flexibility of being able to trade-off power dissipation against performance through power supply reduction gives the designer another degree of freedom in designing circuits. However the reduction in power supply, which also reduces the dynamic power dissipation in the circuit, may have some unintended consequences. Firstly, the supply voltage reduction adversely affects the noise characteristics of the circuit which may result in a circuit malfunctioning. As outlined in Chapter 4, the effects of supply voltage reduction on noise margin and undefined region in a CMOS inverter were investigated; it was found that the

strained-Si implementation has a better overall noise margin than the conventional Si due to the higher mobility and lower threshold voltage which was discussed using the voltage transfer curves (VTC). Secondly, as the supply voltage is reduced to levels close to the threshold voltage, strained-Si circuit was seen to operate better than the conventional Si circuit. This effect was shown by calculating the Minimum Energy Point (MEP) of strained-Si and Si inverters which is defined as the supply voltage point where the active energy and leakage energy cross each other.

Research has shown that the power dissipation in circuits can also be reduced if different transistor configurations are employed, in particular the use of 'stacked' transistors. When strained-Si devices were used in a 'stacked' inverter it was observed that:

- With up to 0.6% strain the power dissipation is similar to that of a conventional un-stacked Si inverter (the power dissipation is seen to rise dramatically beyond this amount of strain).

- Stacked strained-Si inverter (with an amount of strain equal to 0.6%) was 16% faster than that of the non-stacked Si inverter with similar power dissipation.

again highlighting the potential of using strained-Si in low-power high performance circuit applications.

Having investigated the potential advantage of using strained-Si in the realisation of digital circuits the work was subsequently extended, briefly, to consider which of the technologies could be used to improve the performance of analog circuits. Subsequently a CMOS inverting amplifier was designed and a frequency analysis performed; it was found that the gain was higher and the bandwidth of the strained-Si circuit was twice that of the standard CMOS implementation which is due to much higher transcondcutance seen in strained-Si devices.

The work outlined in this thesis has demonstrated that strained-Si has the potential to be used to good advantage as an implementation technology for future

generations of digital and analog circuit applications. However the semiconductor industry, driven by market forces, is not only interested in improving performance and power consumption, but also in increasing functionality per unit area brought about by scaling down device dimensions. Consequently, the final aspect of the strained-Si technology to be investigated is the susceptibility of the device characteristics to process variations. This aspect of the work is considered important as scaling of device dimensions continues, random process fluctuations during manufacture are becoming problematic, transistor behaviour becoming less deterministic and more statistical which subsequently impacts on overall circuit performance rendering it less predictable [3, 4].

The methodology to study the impact of process variability on circuits was explored in Chapter 5. The variability analysis performed was based on a statistical technique called DOE and RSM. This approach was chosen over conventional Monte Carlo statistical technique due to its computational efficiency and being reasonably accurate. The variability analysis was performed in two phases; the first being a screening technique called Plackett-Burman (PB) screening to identify all the parameters in a process which have less impact on circuit output (insignificant). The second phase involves the modeling of the variation of response (output of circuit) based on the variation of significant parameters obtained after screening. To screen out the insignificant parameters, a threshold level is assumed below which it can be considered that a variation in the parameter value has only marginal effect on circuit behavior. The threshold value should be considered with the knowledge about the IC process and this can be considered as a disadvantage due to the fact that the threshold value could be kept at a lower level to include more parameters into the significant group of parameters which makes it computationally inefficient. Another disadvantage with the DOE is that only linear effects between the input parameter and output response are considered in the screening process. However, the interaction effects of the significant parameters with respect to the output response are modeled in the RSM.

The statistical method DOE/RSM was applied to the Jamb Latch and Mutual Exclusion Element (MUTEX) circuits which are basic building blocks used in most asynchronous circuits. The state of the art 65nm technology parameters

were considered to create the transistors. The oxide thickness, substrate doping and the amount of the strain applied in the channel were determined as the process parameters which impact significantly from 17 process parameters (shown in Appendix B) on an important design parameter namely, the metastability resolution time (t_m) of MUTEX. The RSM was then performed for the value of t_m based on these three process parameters. The variation of bi-axial strain in the channel was found to be the most significant process parameter which impacts on the value of t_m.

It was also observed that the effects of the variation in substrate doping (N_{sub}) on the value of t_m was less pronounced in circuits with devices having a higher percentage of strain applied to the channel. Furthermore, the metastability resolution time was also reduced.

Even though the technique of bi-axially straining the channel of transistor to improve the transport properties of transistors is well known in technology community there is a need to collaborate with designers with specific details of technology to improve the yield of circuits. The present design research is limited by the fact that the EDA tools are not updated with latest technology libraries which effectively extends the time to market a product. The methods and techniques used in this thesis could be extended to new technology or design techniques to reduce the time to market a product.

6.2 Conclusions

Market demands for systems with high performance low power dissipation and increased functionality have been the main driving force for technology development by the semiconductor industry. To date planar bulk CMOS has been the main workhorse by which present day and near future systems have been and will be implemented. However standard CMOS technology is coming close to its useful working life for advanced applications, although some additional mileage can be achieved through device scaling, however an increasing number of issues [23] need to be resolved with diminishing returns.

The work in this thesis investigated strained-Si technology as a potential successor to planar CMOS in some high performance applications. The advantages offered by the technology are:

- Indicated an increase in the value of saturated current (21%) which gives the device a better drive capability in logic applications.
- Strained-Si devices could operate with lower supply voltages. This infers a reduction in dynamic power dissipation in circuits.
- With reduced supply voltages, noise characteristics of circuits were not significantly degraded.
- For similar power dissipation as in a conventional CMOS circuit, an improved performance level was achieved in the strained-Si circuit (permits a trade-off between performance and power dissipation).
- Designer has the flexibility to use percentage strain in the device channel to tailor V/I characteristics for use in particular applications.

Overall when strained-Si technology was used in several sample circuit applications listed below and compared with standard CMOS implementation it is found that:

- Ring oscillator performance was 17% faster.
- Inverting amplifier had twice the bandwidth with increased gain.
- MUTEX- metastability resolution time (t_m) was significantly reduced.

These advantages can be achieved without any major changes being required in the basic planar bulk CMOS process; this also includes the ability to vary percentage strain permitting the process to be tailored to specific applications.

Although the strained-Si has the advantages outlined above it has a problem of self-heating which can affect the threshold voltage, subthreshold leakage current and carrier mobility, however these effects can be reduced by introducing a thin virtual substrate.

6.3 Future Work

The continual drive by market forces for faster, lower power and higher complexity systems is creating a spectrum of issues which must be addressed. At one time technology and design issues could be segregated since device behaviour was largely deterministic. However, as device dimensions are reduced device behaviour becomes more statistical which adversely impacts on system design, when, for example, performance and reliability become less predictable. Furthermore, even for a given technology node, one size fits all concept is no longer applicable, that is device characteristics differ depending upon whether it is a low dynamic power, high performance or low stand by power application.

The work in this thesis demonstrated that strained-Si technology considered has the flexibility to be readily used in either high performance or low dynamic power applications (speed/power trade-off). However it does have the disadvantage of incurring a higher stand-by (static) power due to the devices having higher subthreshold currents. Consequently techniques to reduce leakage current require further investigation. Furthermore it is recognised that as gate lengths are shortened the advantage of higher carrier mobility, brought about by strain in the channel is also decreased. Consequently, other methods of increasing carrier mobility need to be investigated assuming that any new technology is compatible with standard CMOS process. The alternative is to investigate entirely new device structures.

In addition to improving the overall performance of systems, advances in the semiconductor technology have also been driven by the requirement to increase functional density by scaling down device dimensions, this has created two areas for further research.

- As device dimensions are reduced present day process/device models are too simplistic to accurately predict process/device behaviour. Consequently new process/device models need to be developed together with a move away for 2D to 3D simulation of device structures.
- Furthermore as device dimensions are reduced random process variations are having a profound effect on device behaviour and the subsequent yield of manufactured circuits. Consequently it is essential to be able to analyse the

effect of process variations and their impact on various aspects of overall circuit behaviour. Unfortunately there are a vast number of input (process) parameters that can vary. The current approach to analysing the effects of random variables is to use Monte Carlo Analysis; unfortunately this approach is inefficient when the number of variables is large. Hence more efficient tools to perform the analysis are required [5-7].

References

[1] International Technology Roadmap for Semiconductors, “Process Integration, Devices and Structures”, 2007 Ed.

[2] MEDICI user manual: *Two-Dimensional Device simulation program.* Sunnyvale: Synopsys, Inc., 2000.

[3] S. Y. Chou, P. R. Krauss, and P. J. Renstrom, "Imprint Lithography with 25-Nanometer Resolution", *Science*, vol. 272, pp. 85-87, 1996.

[4] K. Bernstien, D. J. Frank, A. E. Gattiker, W. Haensch, B. L. Ji, et.al., “High-performance CMOS variability in the 65nm regime and beyond”, *IBM Journal for Research and Development*, vol. 50, pp. 433-459, 2006.

[5] C. Jacoboni, and P. Lugli, *The Monte Carlo method for semiconductor device simulation.* New York: Springer, 1989.

[6] P. N. Koch, D. Mavris, and F. Mistree, "Multi-level, Partitioned Response Surfaces for Modeling Complex Systems," *American Institute of Aeronautics and Astronautics*, pp. 1-15, 1998.

[7] S. Shedabale, H. Ramakrishnan, G. Russell and A. Yakovlev, " Statistical Modelling of Variation for Advanced Process Technologies using Multi-level Partitioned Response Surface Approach”, *Journal of IET Circuits, Devices and Systems* (Accepted for publication).

Appendix A

Parameter Extraction and PSPICE Models

Parameter extraction is performed using AURORA simulator available with Synopsys TCAD suite. AURORA is a device characterisation and parameter extractor system to extract compact circuit model parameters from experimental or simulation device characteristics. Some of the BSIM Level 3 model parameters for the 65nm n-MOS device using AURORA which are used in PSPICE for circuit simulation are shown in A. 1.

```
*$
*BeginParam
*LEVEL=3 (1, 4,1)
*L=0.0650E-6
*W=.5
*KP=40.000E-6
*RS=10.000E-3
*RD=10.000E-3
*VTO=0.2
*RDS=1.0000E6
*TOX=0.00110E-6
*CGSO=40.000E-12
*CGDO=10.000E-12
*CBD=1.0000E-9
*MJ=.5
*PB=.8
*FC=.5
*RG=5
*RB=1.0000E-3
*PHI=.6
*XJ=0.0200e-6
*UO=1000
*$
```

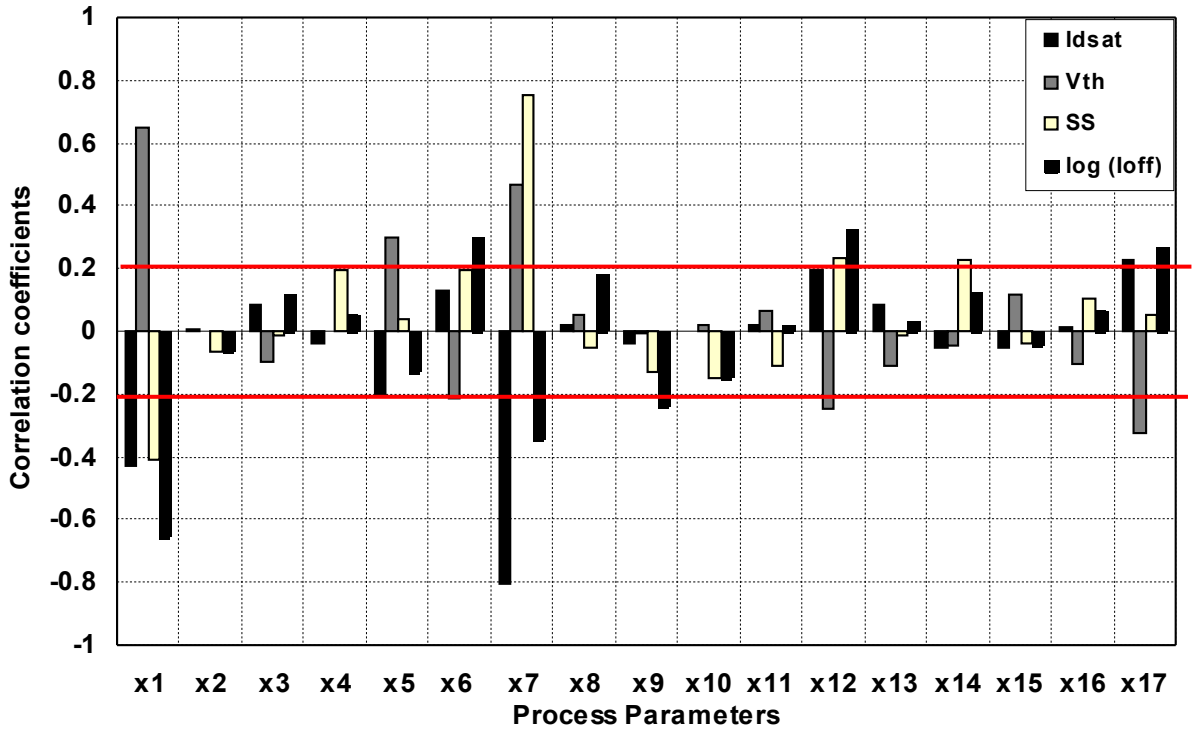
A. 1 Typical Level 3 BSIM parameters used in PSPICE simulations

Appendix B

Parameters for 65nmTechnology Node

Symbol	Description	Unit	Mean Value	Lower level value (-1)	Higher le value (+1
x_1	gate length	nm	25	22.5	27.5
x_2	P – well implant dose	atoms/cm^2	1x10^{12}	0.9x10^{12}	1.1x10^{12}
x_3	P – well implant energy	keV	40	36	44
x_4	P – well diffusion temp.	^{o}C	1000	990	1010
x_5	V_{th} implant dose	atoms/cm^2	4x10^{13}	3.6x10^{13}	4.4x10^{13}
x_6	V_{th} implant energy	keV	8.5	7.65	8.85
x_7	gate – oxide diffusion temp.	^{o}C	800	790	810
x_8	poly Si. Thickness	nm	60	54	66
x_9	pocket halo dose	atoms/cm^2	8x10^{12}	7.2x10^{12}	8.8x10^{12}
x_{10}	pocket halo energy	keV	11	9.9	12.1
x_{11}	pocket halo tilt angle	^{o}	35	31.5	38.5
x_{12}	LDD implant dose	atoms/cm^2	1.4x10^{14}	1.26x10^{14}	1.54x10^{14}
x_{13}	LDD implant energy	keV	6	5.4	6.6
x_{14}	side–wall spacer thickness	nm	30	27	33
x_{15}	S/D implant dose	atoms/cm^2	1.2x10^{15}	1.08x10^{15}	1.32x10^{15}
x_{16}	S/D implant energy	keV	10	9	11
x_{17}	RTA temp.	^{o}C	1000	990	1010

B. 1 Process parameters of a typical 65nm technology node.



B.2 Screening of input process parameters for different device responses.

The horizontal red line shows the threshold value chosen which is 0.2 in this case which can be varied which increases or decreases the design complexity and computation. X1to X17 shown in B.2 is explained in B.1.

Appendix C

C.1 Medici Input deck for 65nm device

```
$***************************65nm State of the art device*****************************
 $About          : 65nm technology device with ITRS requirements, All the parameters are
 $               : derived from MASTAR4 and ITRS, The three files are for HP, LOP and LSTP
 $               : as the name suggests for the same technology node. The effective channel
 $               : lengths are 37, 53, and 65 for HP, LOP, and LSTP respectively with changes
 $               : in other parameters. There is  0.7X, 0.4X, 0.5X rule applied for most of
 $               : of the parameters according to ITRS
 $Version        : 1.0
 $Revision       : 0.1
 $Author         : Hiran
 $File Name      : si_nmos_hp.inp
 $Old file name  : ssi_nmos_idvd_halo2.inp (/home/a3706807/oct05/SOI/IDVD/)
 $Date    : 24-04-2007
 $Date Modified  : First copy of the file
 $Purpose        : Self heating, but scaling the technology node and compare Ioffs (Mirca
sean)
 $Edit summary   : has used the second type of halo used in the directory oct05
 $Important note : NA
 $Models used    : SRFMOB2 FLDMOB ANALYTIC BGN SRH (drift diffusion with phonon
scattering)
 $Hydrodynamic   : SRFMOB2 TMPMOB ANALYTIC II.TEMP/ELE.TEMP in symbolic
(Hydro dynamic models)
 $References     : Medici manual for models
 $Result         : Device used in circuit analysis

 $*************************************************************************************$
 TITLE     0.065 Micron N-Channel MOSFET (IDs Vs VDs)
 $********************* MESH ***********************************************************$
```

```
$Rectangular Mesh

MESH      SMOOTH=1

$SPECIFY THE MESH

x.mesh loc=-0.1500  spac=0.00300

x.mesh loc=-0.1400  spac=0.00300

x.mesh loc= 0.1400  spac=0.00300

x.mesh loc= 0.1500  spac=0.00300

y.mesh loc=-0.0025  spac=0.00050

y.mesh loc=-0.0010  spac=0.00050

y.mesh loc= 0.0010  spac=0.00050

y.mesh loc= 0.0050  spac=0.00090

y.mesh loc= 0.0100  spac=0.00200

y.mesh loc= 0.1000  spac=0.05000


$************************ MESH REFINE ***************************************************$


$Eliminate some unnecessary substrate nodes
ELIMIN    columns  y.min=0.06
ELIMIN    columns  y.min=0.07
ELIMIN    columns  y.min=0.08
ELIMIN    columns  y.min=0.09

$************************ REGIONS *******************************************************$


$ Specify oxide and silicon regions
REGION   name=silicon SILICON x.min=-0.150 x.max= 0.150 y.min= 0.000 y.max= 0.009
REGION   name=SiGe    SIGE    x.min=-0.150 x.max= 0.150 y.min= 0.009 y.max= 0.100
x.mole=0.25
REGION   name=oxide   OXIDE   y.max= 0.000


$***************************ELECTRODES*****************************************************$

$
$Electrode definition
ELECTR    name=Gate    x.min=-0.0325 x.max= 0.0325  y.max=-0.0009
ELECTR    name=Source  x.min=-0.1500 x.max=-0.1200  y.max= 0.0000
ELECTR    name=Drain   x.min= 0.1200 x.max= 0.1500  y.max= 0.0000
ELECTR    name=Substrate  bottom
ELECTR    name=Sink       bottom thermal


$************************** DOPING ******************************************************$
```

```
$Specify impurity profiles and fixed charge
PROFILE   p-type n.peak=8.697e18  uniform out.file=NMOS0.09
PROFILE   p-type n.peak=8.697e18  y.char=0.006

$******************** SOURCE/ DRAIN/ GATE PROFILE *********************************$


$0.09 micron channel derive from 200nm self heating analysed devices


PROFILE   n-type n.peak=1E20 y.junc=0.0140 x.min=-0.1500 width=0.12850
+         xy.rat=0.75


PROFILE   n-type n.peak=3.1E20 y.junc=0.0280 x.min=-0.1500 width=0.06425
+         xy.rat=0.50


PROFILE   n-type n.peak=1E20 y.junc=0.0140 x.min= 0.0215 x.max=0.15000
+         xy.rat=0.75


PROFILE   n-type n.peak=3.1E20 y.junc=0.0280 x.min= 0.08575 x.max=0.15000
+         xy.rat=0.50


$*********************************MATERIALS***********************************************$

material silicon
+         permitti     = 11.900
+         eg300        = 1.04519
+         affinity     = 04.218
+         nc300        = 01e19
+         nv300        = 06e18
+         taun0        = 03e-5
+         taup0        = 01e-5

material sige
+         permitti     = 13.2000
+         eg300        = 1.0612
+         affinity     = 04.0500


$mobility 1153*5 25%Germanium
mobility silicon
+mun0                 = 2133.05

$***************************** INTERFACE PROPERTIES ***************************************$


 $CAN SPECIFY REGIONS;IF NOT SPECIFIED THEN ITS THE INTERFACE BETWEEN
SEMICONDUCTOR AND
 $INSULATOR
 INTERFAC  qf=1E10

$******************************** PLOTS ************************************************$


$see the difference for the following plots
```

```
PLOT.2D    GRID TITLE="n-MOSFET with LDD"  FILL  SCALE

PLOT.3D    DOPING TITLE="DOPING" FILL

SURFACE    C.AUTO

PLOT.1D    DOPING TITLE="DOPING"


$********************************* REGRIDS ***********************************************$


$Regrid on doping
REGRID    DOPING LOG  IGNORE=OXIDE RATIO=3 SMOOTH=2
+         IN.FILE=NMOS0.09

PLOT.2D   GRID  TITLE="Grid after doping"  FILL  SCALE


$********************************* CONTACTS **********************************************$


$Specify contact parameters
CONTACT   NAME=Gate  N.POLY WORKFUNC=4.25


$****************************** CONTACT RESISTANCE ****************************************$


CONTACT   NAME=Sink r.therma=2.9e5


 $********************************* MODELS ************************************************$


$The solution file is taken from the example mdes1.inp
$HOL.TEMP/ELE.TEMP should be used with the TMPMOB in the SYMB statement
$ANALYTIC includes Temperature dependent
$SRFMOB
$FLDMOB
$Drift diffusion model

MODELS SRFMOB2 FLDMOB AUGER BGN ANALYTIC PRINT


$********************************* VGS=10V ***********************************************$

$0-carrier solve with 10v on the gate
SYMB      CARRIERS=0
METHOD    DAMPED  ICCG
SOLVE     V(Gate)=1.2 OUT.FILE=SI_LT5
SYMB      CARR=2  NEWTON  LAT.TEMP COUP.LAT
METHOD    ^AUTONR
$Ramp the drain to 6v
  LOG      OUT.FILE=ssi_latst_wlt
  SOLVE    V(Drain)=0.1  VSTEP=0.05  ELEC=Drain  NSTEP=25

$****************************** CURRENT VOLTAGE PLOTS (VGS=1.2V)
*****************************$
```

```
PLOT.1D   IN.FILE=ssi_latst_wlt X.AXIS=V(Drain)  Y.AXIS=I(Drain) out.file=dd_sf_lt
+         COLOR=5

stop
```

C.2 Medici Input deck for 90 nm device

```
$*****************************************************************************

 $Version        : 1.0.0

 $Author         : Hiran

 $File Name       : ssi_nmos_90nm.inp

 $Old file name  : initial_mesh_ssi_nmos_idvd_halo2.inp
(/home/a3706807/oct05/SOI/IDVD/)/home/a3706807/sept06/90)

 $Revision         : 0.0

 $Date    : 17-07-2006

 $Date Modified  : 01-18-2007

 $About          : silicon NMOS MOSFET for Ids Vs Vds Curve for ULIS

 $Edit summary    : has used the second type of halo used in the directory oct05

 $Important note : see the change when inc the oxide thickness (y.mesh=-0.0090
spac=0.0030)

 $Models used     : SRFMOB2 FLDMOB ANALYTIC BGN SRH (drift diffusion with phonon
scattering)

 $Hydrodynamic   : SRFMOB2 TMPMOB ANALYTIC II.TEMP/ELE.TEMP in symbolic
(Hydro dynamic models)

 $References     : Medici manual for models

 $Result         : 90nm Device used in circuit analysis


 $*****************************************************************************$

 TITLE     0.09 Micron N-Channel MOSFET (IDs Vs VDs)

 $********************* MESH ************************************************$


 $Rectangular Mesh

 MESH      SMOOTH=1

 $SPECIFY THE MESH
```

```
x.mesh loc=-0.2000  spac=0.01000

x.mesh loc=-0.1700  spac=0.00300

x.mesh loc=-0.1400  spac=0.00300

x.mesh loc= 0.1400  spac=0.00300

x.mesh loc= 0.1700  spac=0.00300

x.mesh loc= 0.2000  spac=0.01000

y.mesh loc=-0.0038  spac=0.00050

y.mesh loc=-0.0010  spac=0.00050

y.mesh loc= 0.0010  spac=0.00050

y.mesh loc= 0.0050  spac=0.00050

y.mesh loc= 0.0100  spac=0.00200

y.mesh loc= 0.1000  spac=0.05000


$************************ MESH REFINE ***********************************************$


$Eliminate some unnecessary substrate nodes
ELIMIN    columns  y.min=0.06
ELIMIN    columns  y.min=0.07
ELIMIN    columns  y.min=0.08
ELIMIN    columns  y.min=0.09

$************************ REGIONS ***************************************************$


$ Specify oxide and silicon regions
REGION   name=silicon SILICON x.min=-0.200 x.max= 0.200 y.min= 0.000 y.max= 0.009
REGION   name=SiGe    SIGE    x.min=-0.200 x.max= 0.200 y.min= 0.009 y.max= 0.100
x.mole=0.25
REGION   name=oxide   OXIDE   y.max= 0.000


$***************************ELECTRODES*************************************************$


$Electrode definition
ELECTR    name=Gate    x.min=-0.120 x.max= 0.120  TOP
ELECTR    name=Source  x.min=-0.200 x.max=-0.150  y.max= 0.000
ELECTR    name=Drain   x.min= 0.150 x.max= 0.200  y.max= 0.000
ELECTR    name=Substrate  bottom


$************************** DOPING ***************************************************$


$Specify impurity profiles and fixed charge
PROFILE   p-type n.peak=1.7E17  uniform out.file=NMOS0.09
```

```
 PROFILE   p-type n.peak=1.7E16  y.char=0.006

 $************************** SOURCE/ DRAIN/ GATE PROFILE
**********************************$


 $0.13micron channel derive from 200nm self heating analysed devices


 PROFILE   n-type n.peak=1E18 y.junc=0.020 x.min=-0.2000 width=0.1400
 +        xy.rat=0.75


 PROFILE   n-type n.peak=1E20 y.junc=0.040 x.min=-0.2000 width=0.0700
 +        xy.rat=0.50


 PROFILE   n-type n.peak=1E18 y.junc=0.020 x.min= 0.0600 x.max=0.2000
 +        xy.rat=0.75


 PROFILE   n-type n.peak=1E20 y.junc=0.040 x.min= 0.1300 x.max=0.2000
 +        xy.rat=0.50


material silicon
 +        permitti     = 11.900
 +        eg300        = 01.019282
 +        affinity     = 04.249
 +        nc300        = 01e19
 +        nv300        = 06e18
 +        taun0        = 03e-5
 +        taup0        = 01e-5

 material sige
 +        permitti     = 13.2000
 +        eg300        = 01.0441
 +        affinity     = 04.0500


 $mobility 1153*5
 mobility silicon
 +mun0               = 2133.05


 $***************************** INTERFACE PROPERTIES ***************************************$


 $CAN SPECIFY REGIONS;IF NOT SPECIFIED THEN ITS THE INTERFACE BETWEEN
SEMICONDUCTOR AND
 $INSULATOR
 INTERFAC  qf=1E10

 $******************************** PLOTS ************************************************$


 $see the difference for the following plots
 PLOT.2D    GRID TITLE="n-MOSFET with LDD"  FILL  SCALE
 PLOT.3D    DOPING TITLE="DOPING" FILL
 SURFACE    C.AUTO
```

```
PLOT.1D    DOPING TITLE="DOPING"


$********************************* REGRIDS ***********************************************$


$Regrid on doping
REGRID    DOPING LOG  IGNORE=OXIDE RATIO=3 SMOOTH=2
+         IN.FILE=NMOS0.09
PLOT.2D   GRID  TITLE="Grid after doping"  FILL  SCALE


$********************************* CONTACTS **********************************************$


$Specify contact parameters
CONTACT   NAME=Gate  N.POLY WORKFUNC=4.25


 $********************************* MODELS ************************************************$

COMMENT   Specify physical models to use
MODELS    prpmob  fldmob bgn srh temp=300 print

COMMENT   Symbolic factorization, solve, regrid on potential
SYMB      CARRIERS=0

METHOD    ICCG  DAMPED

SOLVE

REGRID    POTEN  IGNORE=OXIDE  RATIO=.2  MAX=1  SMOOTH=1
+         IN.FILE=NMOS0.09
+         OUT.FILE=NMOS1

PLOT.2D   GRID  TITLE="Example Potential Regrid"  FILL  SCALE

$SAVE MESH OUT.FILE=ssi_nmos W.MODELS


COMMENT   Solve using the refined grid, save solution for later use
SYMB      CARRIERS=0
SOLVE     OUT.FILE=NMOS2

LOAD IN.FILE=NMOS2

symbolic newton carriers=1 electrons
Solve v(gate)=-0.05 v(drain)=0.1

log out.file=si_nmos1

solve electrod=gate vstep=0.05 NSTEP=25

EXTRACT MOS.PARA

plot.1d  in.file=si_nmos1 x.axis=v(gate) y.axis=i(drain) color=1 out.file=ssi_sample

SAVE MESH OUT.FILE=NMOS W.MODELS

STOP
```

C.3 Medici Input deck for 130nm device

```
$*****************************************************************************

 $Version        : 1.0.1

 $Author         : Hiran

 $File Name      : si_nmos.inp

 $Old file name  : ssi_nmos_idvd_halo2.inp (/home/a3706807/oct05/SOI/IDVD/)

 $Revision       : 0.1

 $Date   : 17-07-2006

 $Date Modified  : 17-07-2006

 $About          : silicon NMOS MOSFET for Ids Vs Vds Curve for banglore conference

 $Purpose        : VLSI design conference in balgalore, Koushik

 $Edit summary   : has used the second type of halo used in the directory oct05

 $Important note : see the change when inc the oxide thickness (y.mesh=-0.0090
spac=0.0030)

 $Models used    : SRFMOB2 FLDMOB ANALYTIC BGN SRH (drift diffusion with phonon
scattering)

 $Hydrodynamic   : SRFMOB2 TMPMOB ANALYTIC II.TEMP/ELE.TEMP in symbolic
(Hydro dynamic models)

 $References     : Medici manual for models

 $Result         : Device used in circuit analysis


 $*****************************************************************************$

 TITLE     0.13 Micron N-Channel MOSFET (IDs Vs VDs)

 $********************* MESH ***********************************************$


 $Rectangular Mesh

 MESH      SMOOTH=1

 $SPECIFY THE MESH

 x.mesh loc=-0.2500  spac=0.05000

 x.mesh loc=-0.1700  spac=0.00300

 x.mesh loc=-0.1400  spac=0.00300
```

```
x.mesh loc= 0.1400  spac=0.00300

x.mesh loc= 0.1700  spac=0.00300

x.mesh loc= 0.2500  spac=0.05000

y.mesh loc=-0.0038  spac=0.00050

y.mesh loc=-0.0010  spac=0.00050

y.mesh loc= 0.0010  spac=0.00050

y.mesh loc= 0.0050  spac=0.00050

y.mesh loc= 0.0100  spac=0.00200

y.mesh loc= 0.1000  spac=0.05000


$************************ MESH REFINE ************************************************$


$Eliminate some unnecessary substrate nodes
ELIMIN    columns  y.min=0.07


$************************ REGIONS ****************************************************$


$ Specify oxide and silicon regions
REGION   name=silicon SILICON x.min=-0.250 x.max= 0.250 y.min= 0.000 y.max= 0.009
REGION   name=SiGe    SIGE    x.min=-0.250 x.max= 0.250 y.min= 0.009 y.max= 0.100
x.mole=0.25
REGION   name=oxide   OXIDE   y.max= 0.000


$***************************ELECTRODES*************************************************$


$Electrode definition
ELECTR    name=Gate    x.min=-0.150 x.max= 0.150  TOP
ELECTR    name=Source  x.min=-0.250 x.max=-0.200  y.max= 0.000
ELECTR    name=Drain   x.min= 0.200 x.max= 0.250  y.max= 0.000
ELECTR    name=Substrate  bottom


$*************************** DOPING ***************************************************$


$Specify impurity profiles and fixed charge
PROFILE   p-type n.peak=1.7E17  uniform out.file=NMOS0.13
PROFILE   p-type n.peak=1.7E17  y.char=0.006

$*************************** SOURCE/ DRAIN/ GATE PROFILE
*********************************$


$0.13micron channel derive from 200nm self heating analysed devices
```

```
PROFILE   n-type n.peak=1E20 y.junc=0.020 x.min=-0.2500 width=0.170677
+         xy.rat=0.75

PROFILE   n-type n.peak=1E18 y.junc=0.040 x.min=-0.2500 width=0.1200
+         xy.rat=0.50

PROFILE   n-type n.peak=1E20 y.junc=0.020 x.min= 0.079323 x.max=0.250
+         xy.rat=0.75

PROFILE   n-type n.peak=1E18 y.junc=0.040 x.min= 0.1300 x.max=0.250
+         xy.rat=0.5



$***************************** INTERFACE PROPERTIES ***************************************$


$CAN SPECIFY REGIONS;IF NOT SPECIFIED THEN ITS THE INTERFACE BETWEEN
SEMICONDUCTOR AND
$INSULATOR
INTERFAC  qf=1E10

$********************************** PLOTS ************************************************$


$see the difference for the following plots
PLOT.2D    GRID TITLE="n-MOSFET with LDD"  FILL  SCALE
PLOT.3D    DOPING TITLE="DOPING" FILL
SURFACE    C.AUTO
PLOT.1D    DOPING TITLE="DOPING"


$********************************** REGRIDS **********************************************$


$Regrid on doping
REGRID    DOPING LOG  IGNORE=OXIDE RATIO=3 SMOOTH=2
+         IN.FILE=NMOS0.13
PLOT.2D   GRID  TITLE="Grid after doping"  FILL  SCALE


$********************************** CONTACTS *********************************************$


$Specify contact parameters
CONTACT   NAME=Gate  N.POLY WORKFUNC=4.25


 $********************************** MODELS ***********************************************$

COMMENT   Specify physical models to use
MODELS    prpmob  fldmob  bgn srh temp=300 print

COMMENT   Symbolic factorization, solve, regrid on potential
SYMB      CARRIERS=0

METHOD    ICCG  DAMPED

SOLVE
```

```
REGRID   POTEN  IGNORE=OXIDE  RATIO=.2  MAX=1  SMOOTH=1
+        IN.FILE=NMOS0.13
+        OUT.FILE=NMOS1

PLOT.2D   GRID  TITLE="Example Potential Regrid"  FILL  SCALE

SAVE MESH OUT.FILE=ssi_nmos W.MODELS


stop


COMMENT   Solve using the refined grid, save solution for later use
SYMB      CARRIERS=0
SOLVE     OUT.FILE=NMOS2

LOAD IN.FILE=NMOS2

symbolic newton carriers=1 electrons

solve v=-0.05 v(drain)=2.5


log out.file=si_nmos1

solve electrod=gate vstep=0.05 NSTEP=51

EXTRACT MOS.PARA

plot.1d  in.file=si_nmos1 x.axis=v(gate) y.axis=i(drain) color=1 out.file=si_sample

SAVE MESH OUT.FILE=NMOS W.MODELS

STOP
```

C. 4 Medici Input deck for 400nm device

```
$******************400nm Device deck*****************************************

$MUN.MAX=1429 for analytic default take 1.8 times as value of enhancement

$*******************************************************************************

$Version        : 1.0.1

$Author         : Hiran

$File Name       : opt_newsol1.inp

$Old file name  : opt_solution3.inp

$Revision         : 0.1

$Date    : 03-03-2006

$Date Modified  : 03-03-2006

$About          : strained silicon n-MOSFET for Ids Vs Vds Curve
```

```
$Purpose       : Optimising the device for different Gate voltages

$History       : Changed the structure from the previous file; mentioned in Log book

$Result        : File to use for self heating analysis for strained silicon MOS device

$Models        : Drift diffusion approximation with phonon scattering

$References    : 1. Medici users manual

$                : 2. Parkers model for strain

$                : 3. Glasgow model

$                : 4. Lattice temperatue

$                : 5. 2174 batch comparison/ calibration

$Comment       : Leff taken as 350 -375nm

$*****************************************************************************$

TITLE     0.4 Lg Micron N-Channel MOSFET (IDs Vs VDs)

$************************** MESH ************************************************$


$Rectangular Mesh

MESH      SMOOTH=1

$SPECIFY THE MESH

x.mesh loc=-0.3500  spac=0.00500
x.mesh loc=-0.1700  spac=0.00300
x.mesh loc=-0.1400  spac=0.00300
x.mesh loc= 0.1400  spac=0.00300
x.mesh loc= 0.1700  spac=0.00300
x.mesh loc= 0.3500  spac=0.00500

y.mesh loc=-0.0066 spac=0.00050
y.mesh loc=-0.0010  spac=0.00050
y.mesh loc= 0.0010  spac=0.00050
y.mesh loc= 0.0100  spac=0.00500
y.mesh loc= 0.1000  spac=0.05000
y.mesh loc= 0.2000  spac=0.05000



$************************ MESH REFINE ******************************************$


$Eliminate some unnecessary substrate nodes
ELIMIN    columns  y.min=0.1500
ELIMIN    columns  y.min=0.0500 x.min=-0.2000 x.max=0.2000
ELIMIN    columns  y.min=0.0100 x.min=-0.2000 x.max=0.2000

$************************ REGIONS **********************************************$
```

```
$ Specify oxide and silicon regions
REGION   name=sige    SIGE    x.min=-0.350 x.max= 0.350 y.min= 0.009 y.max= 0.200
x.mole=0.25
REGION   name=silicon SILICON x.min=-0.350 x.max= 0.350 y.min= 0.000 y.max= 0.009
REGION   name=oxide   OXIDE   y.max= 0.000


$*****************************ELECTRODES*****************************************************$


$Electrode definition (Lg=400nm)
ELECTR    name=Gate    x.min=-0.200 x.max= 0.200  TOP
ELECTR    name=Source  x.min=-0.350 x.max=-0.300  y.max= 0.000
ELECTR    name=Drain   x.min= 0.300 x.max= 0.350  y.max= 0.000
ELECTR    name=Substrate  bottom
ELECTR    name=Sink       bottom thermal


$*************************** DOPING ***************************************************$


$Specify impurity profiles and fixed charge
PROFILE   p-type n.peak=1E17  uniform out.file=nssi1
PROFILE   p-type n.peak=1E17  y.char=0.006

$*************************** SOURCE/ DRAIN/ GATE PROFILE
**********************************$


$0.3micron channel

PROFILE   n-type n.peak=0.5E19 y.junc=0.058 x.min=-0.350 width=0.123
+         xy.rat=0.75

PROFILE   n-type n.peak=2E20 y.junc=0.118 x.min=-0.350 width=0.040
+         xy.rat=0.50

PROFILE   n-type n.peak=0.5E19 y.junc=0.058 x.min=0.227 x.max=0.350
+         xy.rat=0.75

PROFILE   n-type n.peak=2E20 y.junc=0.118 x.min= 0.310 x.max=0.350
+         xy.rat=0.50


$***************************** INTERFACE PROPERTIES ***************************************$


$CAN SPECIFY REGIONS;IF NOT SPECIFIED THEN ITS THE INTERFACE BETWEEN
SEMICONDUCTOR AND
$INSULATOR
INTERFAC  qf=4E11


$*****************************MATERIAL PROPERTIES*****************************************$

$MOBILITY
+REG=silicon
+MUN0=2300
+MUP0=900 PRINT
```

```
 MOBILITY
+REG=silicon
+MUN.MAX=2500
+MUP.MAX=900 PRINT
+GSURFN=0.25

MATERIAL
+REG=silicon
+EG300=1.04
+PERMITTI=11.80
+A.TH.CON=0.03
+EG.MODEL=1 PRINT
$+AFFINITY=4.17

MATERIAL
+REG=sige
+EG300=1.06
+PERMITTI=12.90
+AFFINITY=4.0
+A.TH.CON=1.20 PRINT

$+B.TH.CON=1.56E-3
$+C.TH.CON=1.65E-7
$+D.TH.CON=0.0135


$********************************* PLOTS ************************************************$


$see the difference for the following plots
PLOT.2D    GRID TITLE="n-MOSFET with LDD"  FILL  SCALE
PLOT.3D    DOPING TITLE="DOPING" FILL
SURFACE    C.AUTO
PLOT.1D    DOPING TITLE="DOPING"


$********************************* REGRIDS **********************************************$


$Regrid on doping
REGRID    DOPING LOG IGNORE=OXIDE RATIO=3 SMOOTH=2
+         IN.FILE=nssi1
PLOT.2D   GRID  TITLE="Grid after doping"  FILL  SCALE


 $********************************* CONTACTS INCLUDING RESISTANCE
************************$


$Specify contact parameters
CONTACT   NAME=Gate  N.POLY WORKFUNC=4.21
CONTACT   NAME=SOURCE resist=0.4E2
$WORKFUNC=4.21
CONTACT   NAME=DRAIN  resist=0.4E2
$WORKFUNC=4.21


$****************************** CONTACT RESISTANCE ***************************************$
```

```
 CONTACT   NAME=Sink r.therma=2e5

$********************************** MODELS ***********************************************$

 MODELS     FLDMOB ANALYTIC SRFMOB2 AUGER BGN PRINT
 SYMBOLIC   NEWTON carriers=1 electrons
$ +         lat.temp

 METHOD     cont.rhs n.damp

 $***************************** ID - VD SOLUTION AND PLOTS ********************************$

 COMMENT   Set initial Drain and Gate voltages to 0
 SOLVE     v(gate)=0.0
SOLVE     v(gate)=1.1
 $***************************** VGS=3.16V *********************************************************$

 SOLVE     v(gate)=2.66

 SYMB      CARR=1  NEWTON  LAT.TEMP COUP.LAT
 METHOD     cont.rhs n.damp

 LOG       out.file=opt_sample

 SOLVE     ELEC=Drain V(Drain)=-0.05  VSTEP=0.05    NSTEP=61
 $SOLVE     ELEC=Drain V(Drain)=0.5  VSTEP=0.5    NSTEP=4.5
 $SOLVE     ELEC=Drain V(Drain)=2.75  VSTEP=0.01    NSTEP=30
 $SOLVE     ELEC=Drain V(Drain)=0.20   VSTEP=0.001    NSTEP=40

 $SOLVE     ELEC=Drain V(Drain)=0.4  VSTEP=0.05    NSTEP=12
 $SOLVE     ELEC=Drain V(Drain)=2.00   VSTEP=0.01    NSTEP=100
 $SOLVE    ELEC=Drain V(Drain)=1.60  VSTEP=0.005    NSTEP=040
 $SOLVE    ELEC=Drain V(Drain)=1.80  VSTEP=0.060    NSTEP=020
 $SOLVE    ELEC=Drain V(Drain)=2.40  VSTEP=0.600    NSTEP=006


 PLOT.1D   in.file=opt_sample x.axis=v(drain) y.axis=i(drain)
 +         color=2 out.file=nssi_25Ge_2.66Vgs



 $****************************** PARAMETERS EXTRACTION
*************************************$


 EXTRACT MOS.PARA


 $***************************** SAVING IN TIF FORMAT ***************************************$


 SAVE OUT.FILE=TEST.TIF TIF ALL

 $ SOLVE OUT.FILE=SSI_LATTEMP_HD TIF ALL
```

```
$***************************** BAND DIAGRAM PLOT ****************************************$
 plot.1d val neg x.st=-0.0 x.en=-0.0 y.st=0 y.en=0.1 min=-1 max=1
 plot.1d con neg unchange x.st=-0.0 x.en=-0.0 y.st=0 y.en=0.1
 plot.1d pot neg unchange x.st=-0.0 x.en=-0.0 y.st=0 y.en=0.1
 plot.1d qfp neg unchange x.st=-0.0 x.en=-0.0 y.st=0 y.en=0.1

 $******************************** END PROGRAM ***************************************************$
```