# FABRICATION AND CHARACTERIZATION OF TUNNELING FIELD EFFECT TRANSISTORS (TFETs)

YANG LITAO

NATIONAL UNIVERSITY OF

SINGAPORE

2010

## FABRICATION AND CHARACTERIZATION OF TUNNELING FIELD EFFECT TRANSISTORS (TFETs)

## YANG LITAO

B. Eng (Hons.), NUS

# A THESIS SUBMITTED FOR THE DEGREE OF MASTER OF ENGINEERING

# DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

## NATIONAL UNIVERSITY OF SINGAPORE

2010

#### ACKNOWNLEDGEMENTS

I would like to express my gratitude to my advisor, Dr. Yeo Yee-Chia for giving me this chance to experience the research environment in my period of graduate study in National University of Singapore. He is an admirable academic professional and research pioneer. His sense of responsibility and strict attitude impressed me very much. His advice would benefit me for life.

I would also like thank my co-advisors. I would like to thank Prof. Heng Chun Huat for his insightful suggestions and valuable discussion in my research topic. I would also like to express my thanks to Dr. Patrick Lo, for facilitating my device fabrication work in the Institute of Microelectronics (IME).

I would also like to thank Prof. Samudra, for his kind advice in our project discussions.

Special thanks to my fellow colleagues in Silicon Nano Device Lab, for their valuable discussions, and enjoyable activities we joined together. I also thank the research staffs and engineer assistants at IME, for their kind support of my fabrication work.

Lastly, I would like to take this opportunity to thank my fellow friends and my family, who gave me their wholehearted support during my period of study. Especially, I'd like to thank my wife, for her fully understanding and support to me, especially when I was in the most difficult time. This work would be dedicated to my friends and family.

## **Table of Contents**

Acknowledgements	i
Table of Contents	ii
Abstract	iv
List of Tables	vi
List of Figures	vii
List of Symbols	xi
5	

1
1
5
8
. 12
. 14

Chapter 2. Experimental Study of Tunneling Field Effect Transistors	
2.1 Tunneling FET Process Flow	19
2.2 Key Process Challenges for Tunneling FET Fabrication	
2.2.1 Doping Profile Requirement	
2.2.2 EOT Requirement	
2.2.3 Material System Requirement	
2.2.3.1 Small band gap material	
2.2.3.2 Hetero-junction structures	
2.2.4 Device Body Dimension	30
2.2.5 Alignment Issue	30
2.3 High-κ Metal Gate SiGe Tunneling FET	
2.3.1 Experimental Realization of SiGe Tunneling FET	33
2.3.2 Results and Discussion	35
2.4 Dopant Segregation (DS) Technique and Its Potential	Application
in Tunneling FETs	
2.4.1 Experimental Procedure	44
2.4.2 Results and Discussion	
2.5 Summary	51
Reference	52

Chapter 3. A non-local band to band tunneling algorithm and its application	
in tunneling FET variability study	54
3.1 Development of a New Non-local algorithm for band to band tunneling.	
	54

3.2 Application of the algorithm in Tunneling FET variability study	65
3.2.1 Calibrition of tunneling mass	66
3.2.2 Variability study of a Ge Tunneling FET	71
3.2.2.1 The base structure	75
3.2.2.2 EOT Variation	78
3.2.2.3 Junction position Variation	84
3.3 Summary	91
Reference	92

Chapter 4. Conclusion and future trends of Tunneling FET research	. 94
Reference	. 98

Appendix A: List of Publications

### ABSTRACT

CMOS device scaling faced several fundamental limits as transistor gate length is reduced towards sub-10 nm regime. The conventional Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET)'s subthreshold swing has a limit of 60 mV/decade. This thermal limit has slowed down supply voltage scaling. As a result, power density in modern integrated circuits (ICs) has been increased a lot. New device concepts which can overcome the thermal limit on subthreshold swing have attracted a lot of research interest. Tunneling field effect transistor (TFET or Tunneling FET) is one of the device concepts which can potentially break the 60 mV/decade thermal limit, achieving very abrupt subthreshold swing.

In this study, both experimental and simulation studies of the tunneling FET were carried out.

Important process parameters for fabrication were studied. In order to improve tunneling FET device performance, thin Equivalent Oxide thickness (EOT) with low gate leakage current and abrupt doping profile with degenerated doping concentration should be achieved. High- $\kappa$  metal gate SiGe tunneling FET were fabricated and measured. The potential application of dopant segregation technique in fabrication of tunneling FET was also discussed. Tunneling FET device with dopant segregated source was fabricated and characterized.

A non-local band to band tunneling algorithm was developed. This algorithm was shown to be useful in tunneling FET simulation. Tunneling FET variability study was carried out based on this algorithm. Tunneling FET performance was found to have little dependence on doping concentration variation. However, it was found to be very sensitive to EOT variation and tunneling junction position variation. In order to achieve consistent TFET device performance, very stringent process control is required.

In summary, the TFET device is an attractive device candidate to succeed CMOS technology. However, experimental realization of tunneling FET is very challenging, and is due to the requirement of tight fabrication requirement and stringent process control.

## **List of Tables**

Table 1.1	History	of $V_{dd}$	scaling	and	ITRS	Projection	for	High	Performance	(HP)
MOSFET	s			••••			••••			4

## **List of Figures**

Figure 2.3 A simple process flow of a typical lateral TFET device process......20

Figure 2.7 Plot of simulated  $I_D$ - $V_G$  characteristics for a single-gate tunneling FET with SiGe source. Increasing the Ge content enhances on-state current and subthreshold swing, while off-state leakage is limited by the drain side, the *S* values for the devices

Figure 2.14 Simulated source doping profile along the vertical direction......40

Figure 2.16 Idea for application of dopant segregation in fabrication of tunneling FETs. Shown on the top is a schematic of a DSS tunneling FET. The tunneling junction is zoomed in and shown in the middle, with the corresponding band diagram along the dashed line shown at the bottom. Current conduction is controlled by the Schottky tunneling junction in series with the band to band tunneling junction. The metal to semiconductor contact would behave as an ohmic contact attributed to the elevated doping concentration in the segregated layer. Due to the low band to band

tunneling rate, it is expected to be the current limiting factor. Tunneling FET behavior is expected43
Figure 2.17 Boron (B) profiles of DS junctions using CoSi <sub>2</sub> (left) and NiSi (right), taken by backside SIMS
Figure 2.18 An SEM image of the ultra-thin body tunneling FET with drain side masked45
Figure 2.19 A typical $I_D$ - $V_G$ characteristics of a fabricated device silicided with 5 nm of Ni. The device gate length is around 1 $\mu$ m, device width is around 0.8 $\mu$ m, device drain side is biased at 2.5 V
Figure 2.20 $I_D$ - $V_G$ characteristics of a TFET device silicided by 15 nm of Ni. The device gate length is around 2 $\mu$ m, device width is 0.8 $\mu$ m, device drain side is biased at 2.5 V
Figure 2.21 Atomic resolution Z-contrast images in cross section of an As-doped device, silicide by Co
Figure 3.1 A typical mesh for a gated pin structure for TFET simulation60
Figure 3.2 At a particular energy level, tunneling paths between $E_c$ front and $E_V$ front are found on the mesh. Each tunneling path accounts for the tunneling current in a strip defined between $E_c$ front and $E_V$ front
Figure 3.3 Band Diagram of a tunneling junction visualized in 3D. Tunneling paths at a certain energy level are indicated by black arrows
Figure 3.4 (a) A mesh for a $p^+$ -n 2D tunneling diode, generated by GSS simulator. (b) A manually created mesh for the same diode which is denser than that in (a). (c) <i>I-V</i> characteristics obtained by simulation the two diodes meshed in (a) and (b), using GSS non-local BTBT algorithm
Figure 3.5 Semi-log plots of current-voltage characteristics at 200, 300, and 350K in a Ge p-n diode, where $N_A \sim 2.4 \times 10^{18} \text{ cm}^{-3}$ and $N_D \sim 10^{19} \text{ cm}^{-3}$ 67
Figure 3.6 The 1D diode structure used for Ge tunneling mass fitting
Figure 3.7 The fitting of diode simulation I- V with the experimental data in log-linear (upper) and linear-linear (lower) scale. Simulation data with different tunneling mass is shown on the same figure
Figure 3.8 Graphical depiction of the two definition of sub-threshold swing, point $S$ (calculate from blue dashed line) and average $S$ (calculate from red solid line). It can

be seen that the blue dashed line has a steeper slope, which yields a smaller *S* value, comparing with the average *S* value obtained from the red solid line......74

Figure 3.9 The base structure used for the variability study......75

Figure 3.17 Effect of junction misalignment on the average S value......90

# List of Symbols

$P_{dyn}$	The dynamic power consumption of a transistor
<b>P</b> <sub>standby</sub>	The standby power consumption of a transistor
f	Ttransistor device operation frequency
С	Device load capacitance
W	Device width
Ion	On-state current of a transistor
$I_{off}$	Off-state leakage current of a transistor
$V_{dd}$	Supply voltage of a transistor
S	Subthreshold swing of a transistor
k	Boltzmann constant
Т	Absolute temperature
q	Unit charge
$C_{ox}$	Gate oxide capacitance
$C_d$	Depletion capacitance
Cit	Capacitance associated with interface-trap density
W <sub>tunnel</sub>	Band to band tunneling barrier width
к	Dielectric constant
t	Physical thickness of a dielectric layer
$G_{BTBT}$	Band to band generation rate
$D_{tunnel}$	Band to band tunneling factor
$A_{BTBT}$	Material related parameters for band to band tunneling
B <sub>BTBT</sub>	Material related parameters for band to band tunneling

$C_{BTBT}$	Material related parameters for band to band tunneling
$E_g$	Material band gap
$E_{v,1}, E_{Fp,1}$	Valence band energy and hole Fermi energy on one side of a tunnel barrier
$E_{C,2}, E_{Fn,2}$	Conduction band energy and electron Fermi energy on the other side of a tunneling barrier
Ε	Electric field
$m_r$	Reduced tunneling mass
ħ	Reduced Planck constant
$E_C$	Conduction band energy
$E_V$	Valence band energy
T <sub>tun</sub>	Tunneling probability
V	Tunneling barrier height
Ebar	Energy of a tunneling electron
$n_c$	Electron concentration at the node right behind $E_C$ front
$p_v$	Hole concentration at the node behind the $E_V$ front
D	Area ratio of experimental diode and simulated diode for tunneling mass fitting
$m_0$	Free electron mass
m <sub>lh</sub>	Light-hole mass
$m_c$	Conduction band electron energy
$m_L$	The longitudinal electron mass of the (111) minima
$m_T$	The transverse electron masse of the (111) minima
$V_{th}$	Threshold voltage

$t_{body}$	Body thickness of a SOI wafer
$V_d$	Drain voltage of a transistor
$V_g$	Gate voltage of a transistor
Qchannel	The total channel charge
$t_{ox}$	Dielectric thickness
V <sub>ox</sub>	The potential drop across the gate dielectric of a MOS transistor
L <sub>ov</sub>	Tunneling junction to gate edge misalignment of a tunneling FET

## Chapter 1

## Introduction

#### **1.1 MOSFET scaling in the semiconductor industry**

The semiconductor industry developed rapidly over the past half a century, since the invention of integrated circuits (ICs) in the 1950s. The Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET) is the most important building block of modern high-density IC. MOSFET or device scaling plays an important role in the rapid development of the semiconductor industry. Some early smallscale integrated circuits consist of as few as two transistors each, forming a simple logic gate. Today, an advanced Ultra-Large-Scale-IC (ULSI) such as a Central Processing Unit or a Graphics Processor Unit consists of billions of transistors. Device scaling allows for more devices and/or functions to be integrated into a single chip with a given silicon area, or allows the same number of devices or a given function to be realized on a chip with a smaller silicon area. Cost per device and/or per function has been greatly reduced, which is one of the root reasons for the widespread adoption of electronic devices.

Gordon Moore, co-founder of Intel cooperation, made an observation in 1965 which is now known as the "Moore's Law" [1.1]. It states that the number of transistors being integrated into ICs will increase exponentially, doubling every 2 years. The numbers of MOSFETs integrated into various Central Processing Units (CPUs) followed Moore's Law for the past few decades. The number of transistors in a single IC has increased by more than 6 orders of magnitude over the last 40 years. These numbers are plotted in Figure 1.1(a). The dashed line here shows the prediction by Moore's Law, based on data from the 1970's. Figure 1.1(b) shows the reduction in cost of semiconductor transistors. More than 100 times reduction in the cost per transistor was achieved in the past 20 years. MOSFETs have also been greatly miniaturized. In year 2009, the most advanced technology generation in mass production in the semiconductor has reached the 32 nm technology node, where the gate length is about 25 nm.

Device scaling also leads to shorter gate delay due to an increase in MOSFET drive current with gate length reduction. A shorter gate delay enables a high switching frequency in logic operations. However, as the number of transistors and the operation frequency are increased, the operation power was also greatly increased. The dynamic power  $P_{dyn}$  and standby power  $P_{standby}$  per device can be described by the following equations:

$$P_{dyn} = f \times C \times V_{dd}^{2}, \qquad \qquad \text{Equation 1.1}$$

$$P_{standby} = W \times I_{off} \times V_{dd}, \qquad \qquad \text{Equation 1.2}$$

where *f* is the device operation frequency, *C* is the device load capacitance, *W* is the device width,  $I_{off}$  is the off-state leakage current, and  $V_{dd}$  is the supply voltage.



Figure 1.1 (a) Transistor count in CPUs and prediction by Moore's law. (b) The reduction of cost per transistor in IC manufacturing [1.2].

It is noted that the power consumption for modern ICs needs to be maintained at an acceptable or a sufficiently low level. This is especially important for mobile electronic devices, such as laptop computers, mobile phones and other handheld devices, from both energy consumption and device thermal management considerations. With the number of transistors integrated on a single IC increasing, Equation 1.1 and Equation 1.2 indicate that the supply voltage  $V_{dd}$  has to be scaled down accordingly to achieve this target. Table 1.1 shows the supply voltage in the recent technology nodes (up to 32-nm node) as well as the prediction beyond the 32 nm node from International Technology Roadmap for Semiconductors (ITRS). It can be seen that the voltage scaling rule has not been followed since the 90 nm node.

This recent slowdown in voltage scaling is attributed to one fundamental limit of conventional MOSFETs, i.e. the 60 mV/decade subthreshold swing at room temperature. This limit originates from the fundamental physics of operation of a MOSFET, which will be described in detail in Section 1.2. As a result of this fundamental limit, the power density in modern micro-processor chips has been increasing rapidly with increase in operation frequency. The standby power which is proportional to the off-state leakage current is taking up a larger portion of the total power consumption in modern ICs.

Node (nm)	250	180	130	90	65	45	32	22	14
<i>V<sub>dd</sub></i> (V)	2.5	1.8	1.3	1.2	1.1	1.0	0.9	0.8	0.65

Table 1.1. History of  $V_{dd}$  scaling and ITRS Projection for High Performance (HP) MOSFETs [1.3].

#### **1.2 Conventional MOSFETs and Their Limits**

Figure 1.2 shows a typical n-channel MOSFET (NMOSFET) having heavily doped n-type source and drain regions separated by a p-type channel. In the NMOSFET, the on-state source-to-drain current  $I_{DS}$  is carried by electrons injected from the source to the inversion layer, which flows to the drain (Fig. 1.2, bottom). It should be noted that the energy distribution of electrons in the conduction band in the source follows the Fermi-Dirac distribution.

As the NMOSFET is turned off with a reduction in  $V_{GS}$ , the total number of electrons which can be injected from the source to the channel is reduced exponentially. This leads to an exponential decrease in  $I_{DS}$  with a linear reduction in  $V_{GS}$ . This region of the  $I_{DS}$ - $V_{GS}$  curve is known as the subthreshold region. A parameter called the subthreshold swing S (in units of mV/decade) is defined as the change in  $V_{GS}$  needed to change  $I_{DS}$  by an order of magnitude.

Of the many transistor performance parameters, the subthreshold swing is an important factor that affects  $V_{dd}$  scaling, given the required on-state current  $I_{on}$ and off-state current  $I_{off}$ . A device with a smaller or steeper *S* would have its  $I_{DS}$ modulated over orders of magnitude with a smaller change in  $V_{GS}$ . In a device with a smaller *S*, a lower  $V_{dd}$  would be sufficient for device operation for a given  $I_{on}$  and  $I_{off}$  requirement. This leads to a smaller power consumption when operated in the same operation frequency. Alternatively, for the same  $V_{dd}$ , a device with a smaller *S* would achieve a lower  $I_{off}$  for a given  $I_{on}$ . Both dynamic power and standby power consumption can thus be reduced.



Figure 1.2. A typical n-MOSFET structure is shown at the top. Its energy band diagram along a horizontal line from the source to the drain is illustrated for a device in the off-state (i.e.  $V_{GS} < V_{th}$ , middle) and in the on-state (i.e.  $V_{GS} > V_{th}$ , bottom), for a given  $V_{DS}$ .

The subthreshold swing of a conventional MOSFET can be derived as [1.4]

$$S = \ln 10 \times \left(\frac{kT}{q}\right) \times \left(\frac{C_{ox} + C_d + C_{it}}{C_{ox}}\right),$$
 Equation 1.3

where k is Boltzmann constant, T is the temperature, q is the unit charge,  $C_{ox}$  is the gate oxide capacitance,  $C_d$  is the depletion capacitance, and  $C_{it}$  is the capacitance associated with interface-trap density. In the ideal condition where  $C_{ox}$  dominates over all other capacitances, the S of a conventional MOSFET would be as small as about 60 mV/decade (ln10×kT/q) at room temperature (300 K). This value marks

the smallest *S* achievable for a conventional MOSFET. Thus, *S* of conventional MOSFETs is confined by the thermal limit of kT/q. For a given an on-state leakage current requirement, voltage scaling of 0.2 V would increase the off-state current by more than 3 orders of magnitude. In order to support the voltage scaling requirement and to reduce the power consumption, new device concepts that can overcome the thermal limit of 60 mV/decade are required.

Such a concept is illustrated in Figure 1.3. A transistor with a steeper *S* allows for transition from the on to the off states over a smaller  $V_{GS}$  change, keeping the same  $I_{on}/I_{off}$  current ratio. Such devices are potentially useful in both Low Power (LP) and High Performance (HP) applications.



Figure 1.3. Transistor with steeper S allows for a sharper transition between the on- and off- states, i.e. over a smaller  $V_{GS}$  change. The  $I_{on}/I_{off}$  current ratio is kept the same.

#### **1.3 Device Concepts For Devices With Steep Subthreshold Swing**

Conventional MOSFETs cannot overcome the kT/q thermal limit. Several new device concepts which would allow steep subthreshold swings to be realized have been proposed. The most promising device concepts which has attracted much research attention in recent years are the impact-ionization MOS transistor (I-MOS) and the Tunneling Field Effect Transistor (TFET or tunneling FET).

The I-MOS transistor utilizes the phenomenon of impact ionization to achieve a very steep subthreshold swing. This concept was initially proposed in year 2002 [1.5]. A very abrupt turn-on had been observed through experimental demonstrations, and confirmed or explained with simulations. However, the supply voltage required for operation of such a device remains very high, due to the high electric field required for impact ionization to occur [1.5], [1.6]. Such a requirement restricts its application in modern ICs. Besides the high supply voltage required, there is another major drawback of the I-MOS transistor. As it makes use of the impact-ionization phenomenon, a large number of hot carriers are produced during the device turn-on, and this leads to poor reliability [1.5]. Generally, the poor device reliability hinders the application of the I-MOS transistor.

The Tunneling field effect transistor (TFET) is another device concept that can realize sub-60 mV/decade subthreshold swing. Such a device concept has been proposed as early as the 1990s [1.7]-[1.9]. The basic structure of such a device is a gated p-i-n diode, where the p-i-n region can be oriented either vertically or laterally. A typical lateral structure is as shown in Figure 1.4. In a nchannel tunneling FET (N-TFET), the  $p^+$  source is grounded and the  $n^+$  drain is positively biased. In the off-state where  $V_{GS} = 0$  or  $V_{GS} < V_{th}$ , the tunneling FET resembles a reverse-biased p-i-n diode, and an extremely low off-state current flows. Such a low off-state current has been verified both through simulations [1.10]-[1.13] and experiments [1.9], [1.14].

A positive gate voltage is applied to the gate to turn the N-TFET on. Under this bias condition, a high electric field would be induced near the source to channel junction. The tunneling barrier width  $W_{tunnel}$  between conduction band of the channel side and valence band of the source side would be narrowed down as shown in Figure 1.4. According to quantum mechanics, significant band-to-band tunneling across this barrier could occur when the barrier becomes sufficiently thin, e.g. several nanometers. The valence band electrons from the p<sup>+</sup> doped source could tunnel through the barrier into the conduction band in the channel. This leads to an on-state current flowing between the source and the drain. A schematic showing the N-TFET structure and operation mechanism is depicted in Figure 1.4. In the off-state, the tunneling barrier width  $W_{tunnel}$  is large (typically larger than 5 nm), and negligible band-to-band tunneling is expected; only thermally generated leakage current in the reverse-biased p-i-n diode makes up the off-state current. This leakage current is extremely small.

Due to the band-to-band tunneling nature of the current conduction, the *S* value of a tunneling FET is not subjected to the 60 mV/decade thermal limit like the conventional complementary MOSFETs (CMOSFETs or CMOS).



Figure 1.4. Device structure for a typical N-TFET (top) and the band diagram for such a device in Off (middle) and On (bottom) state. The tunneling barrier width  $W_{tunnel}$  can be modulated by the gate voltage, and turn on or off the device.

Two different device structures or configurations for TFET had been proposed and investigated, each of which has their own advantages and limitations.

The p-i-n structure can be arranged vertically on the substrate [1.15]-[1.18]. Such a structure allows for easier source/channel material engineering using advanced epitaxial growth techniques that are currently available. Hetero-junction TFET had been proposed for device performance enhancement [1.19]-[1.20] by optimal choice of source and channel materials with favorable band alignment for band-to-band tunneling to occur. However, the process flow is not very compatible with conventional CMOS technology, and a new process flow needs to be developed and optimized.

The lateral p-i-n TFET is the other structure which has attracted a lot of research interest. Such a structure can be formed using a layout that is compatible with conventional CMOSFET technology. The additional masking steps for the formation of source/drain with opposite doping types mark the difference between the TFET and the conventional MOSFET. The lateral TFET structure is not self-aligned, and this might affect the ease of integrating nanoscale lateral TFETs .

### **1.4** Objective of research and outline of the thesis

This thesis is aimed at the study of TFET characteristics through both experiments and simulation. In this study, fabrication of lateral TFETs was carried out. Fabricated devices were physically and electrically characterized and analyzed. Besides an experimental study of TFET, a more physical and accurate band-to-band tunneling simulator was also developed. Using this simulator, TFET characteristics were simulated and analyzed, focusing on the impact of process-related variations on device performance. Based on this study, a summary will be given, discussing the future of the lateral TFET and its prospect for application in the semiconductor industry.

The thesis is organized as follows:

In Chapter 1, an introduction on the background of tunneling FET research is given. MOSFET scaling and fundamental limits of conventional CMOSFET are briefly discussed. Novel device concepts that enable the achievement of sub-60 mV/decade subthreshold swing are reviewed.

Chapter 2 documents the experimental findings of this research. Key challenges in the fabrication of tunneling FET with good device performance will be reviewed through a simulation study. A few experimental attempts were made to realize a TFETs. Experimental results will be presented, followed by a discussion.

Chapter 3 focuses on a simulation study of the TFET. A new physical and robust non-local band to band tunneling simulation algorithm was developed and tested on TFET structures. Using this algorithm, TFET performance variation due to process variations was studied. The results will be presented and discussed in this chapter.

Based on the experiment and simulation study of the tunneling FET device presented, a summary will be given in Chapter 4. Recommendations on future study of TFET will be provided.

#### **Reference:**

- [1.1] G. E. Moore, "Cramming more components onto integrated circuits," *Proceedings of the IEEE*, vol. 86, pp. 82-85, 1965.
- [1.2] R. Goodall, D. Fandel, A. Allan, P Landler and H. R. Huff, "Long-Term Productivity Mechanisms of The Semiconductor Industry," *American Electrochemical Society Semiconductor Silicon 2002 Proceedings*, 9<sup>th</sup> edition, May 2002, pp. 125.
- [1.3] International Technology Roadmap for Semiconductors, accessible online at http://www.itrs.net/.
- [1.4] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3 ed.: John Wiley & Sons, INC., 2006.
- [1.5] K. Gopalakrishnan, P. B. Griffin and J. D. Plummer "I-MOS: A novel semiconductor device with a subthreshold slope lower than kT/q," San Francisco, CA, United states, 2002, pp. 289-292.
- [1.6] E.-H. Toh, G. H. Wang, L. Chan, G. Samudra and Y.-C. Yeo, "A double-spacer I-MOS transistor with shallow source junction and lightly doped drain for reduced operating voltage and enhanced device performance," *IEEE Electron Device Letters*, vol. 29, pp. 189-91, 2008.
- [1.7] T. Baba, "Proposal for surface tunnel transistors," *Japanese Journal of Applied Physics, Part 2 (Letters)*, vol. 31, pp. 455-7, 1992.
- [1.8] T. Uemura and T. Baba, "Characterization of depletion-type surface tunnel transistors," *Japanese Journal of Applied Physics, Part 2 (Letters)*, vol. 31, pp. 1727-9, 1992.
- [1.9] W. M. Reddick and G. A. J. Amaratunga, "Silicon surface tunnel transistor," *Applied Physics Letters*, vol. 67, pp. 494-6, 1995.
- [1.10] K. Boucart and A. M. Ionescu, "Double-gate tunnel FET with high-k gate dielectric," *IEEE Transactions on Electron Devices*, vol. 54, pp. 1725-33, 2007.
- [1.11] E.-H. Toh, G. H. Wang, L.Chan, G. Samudra and Y.-C. Yeo, "Device physics and guiding principles for the design of double-gate tunneling field effect transistor with silicon-germanium source heterojunction," *Applied Physics Letters*, vol. 91, pp. 243505-1, 2007.
- [1.12] E.-H. Toh, G. H. Wang, L.Chan, G. Samudra and Y.-C. Yeo, "Device design and scalability of a double-gate tunneling field-effect transistor with silicon-germanium source," *Japanese Journal of Applied Physics*, *Part 1: Regular Papers and Short Notes and Review Papers*, vol. 47, pp. 2593-2597, 2008.
- [1.13] E.-H. Toh, G. H. Wang, L. Chan, G. Samudra and Y.-C. Yeo, "Device physics and design of double-gate tunneling field-effect transistor by silicon film thickness optimization," *Applied Physics Letters*, vol. 90, 2007.
- [1.14] K. K. Bhuwalka, et al., "P-channel tunnel field-effect transistors down to sub-50 nm channel lengths," Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers, vol. 45, pp. 3106-3109, 2006.
- [1.15] K. K. Bhuwalka, J. Schulze and I. Eisele, "Vertical tunnel field-effect transistor with bandgap modulation and workfunction engineering," Leuven, Belgium, 2004, pp. 241-244.

- [1.16] K. K. Bhuwalka, J. Schulze and I. Eisele, "Scaling the vertical tunnel FET with tunnel bandgap modulation and gate workfunction engineering," *IEEE Transactions on Electron Devices*, vol. 52, pp. 909-17, 2005.
- [1.17] K. K. Bhuwalka, J. Schulze and I. Eisele, "Performance enhancement of vertical tunnel field-effect transistor with SiGe in the p+ layer," *Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers*, vol. 43, pp. 4073-4078, 2004.
- [1.18] P.-F. Wang, T. Nirschl, D. Schmitt-Landsiedel and W. Hansch, "Simulation of the Esaki-tunneling FET," *Solid-State Electronics*, vol. 47, pp. 1187-1192, 2003.
- [1.19] A. S. Verhulst, *et al.*, "Complementary silicon-based heterostructure tunnel-FETs with high tunnel rates," *IEEE Electron Device Letters*, vol. 29, pp. 1398-401, 2008.
- [1.20] A. S. Verhulst, W. G. Vandenberghe, K. Maex and G. Groeseneken, "Boosting the on-current of a n -channel nanowire tunnel field-effect transistor by source material optimization," *Journal of Applied Physics*, vol. 104, 2008.

## Chapter 2

# Experimental Study of Tunneling Field Effect Transistors

This chapter documents the experimental study on the fabrication of tunneling field effect transistors (TFETs or tunneling FETs), performed in the Silicon Nano Device Lab (SNDL) and the Institute of Microelectronics (IME). The fabrication of TFETs involves integration of advanced process modules such as metal-gate, high-k dielectric formation. Key process requirements for successful fabrication of tunneling FETs will be highlighted, supported with simulation results. The integration process will also be presented and discussed.

The TFET possesses good performance characteristics in terms of low offstate leakage current and steep subthreshold swing (*S*), as discussed in Chapter 1. Although simulation shows that sub-60 mV/decade swing can be easily achieved with a well-designed device structure, experimental realization is very challenging. Subthreshold swing values of up to a few hundred mV per decade have been reported in several experimental studies [2.1], [2.2].

Two different TFET structures have attracted much research interest in experimental studies in the past few years. The two structures differ from each other in the orientation of how the gated p-i-n diode is being laid on the substrate.



Figure 2.1. Schematic of a typical vertical tunneling FET (left) and a lateral tunneling FET (right).

The p-i-n diode can either be laid vertically or laterally, as shown in Figure 2.1. Each of the two structures has its own advantages and intrinsic drawbacks.

The vertical structure allows easy integration of semiconductor heterostructures designed for optimal band-to-band tunneling. Studies were carried out to investigate the potential of using Ge/Si or InGaAs/Si [2.3], [2.4] heterotunneling junction in TFETs to boost the drive current and to reduce the subthreshold swing. Advanced material growth techniques would greatly help in realizing such hetero-structures in a vertically stacked p-i-n structure. However, the process for making a vertical tunneling FET is not easily integrated with a conventional CMOS process. A new process flow for forming a vertical tunneling FET structure together with CMOS needs to be developed and optimized. It is anticipated that TFET may replace the CMOS core logic devices, while the input/output (I/O) devices in an integrated circuit may still comprise CMOS devices.

The lateral tunneling FET structure is more process-compatible with the conventional CMOS process flow. In fact, the tunneling FET process flow can be

modified from the standard MOSFET process flow by addition of two more mask layers. Each of the two mask layers either masks the source or the drain region, such that different doping types can be introduced into the source and drain regions.

In view of the process-compatibility with conventional CMOS, the lateral tunneling FET structure was chosen for device fabrication in this work.

### 2.1 Tunneling FET Process Flow

The lateral TFET structure was selected for its process compatibility with conventional MOSFET process flow. Two addition mask layers were added to an existing MOSFET fabrication flow, for performing  $p^+$  source and  $n^+$  drain (*vice versa*) implants. The two mask layers are depicted in Figure 2.2.

In the next few paragraphs, a process flow that can be used for fabricating a lateral tunneling FET will be described. Key steps in the process flow are listed in Figure 2.3.

Starting from a silicon wafer, active area definition would be carried out, forming active regions similar to that in a conventional MOSFET process.



Figure 2.2. Two additional masks are added to an existing MOSFET process flow, masking out the source or drain regions for separate implantations. Shown here is the top view of the mask superimposed on a tunneling FET. The cross-hatched regions are opaque and the white regions are transparent. The mask on the left is used during source implant, and the mask on the right is used during drain implant.



Figure 2.3. A simple process flow of a typical lateral TFET device process.

Gate stack (gate electrode on gate dielectric) would be formed. A layer of hard mask (SiO<sub>2</sub>) would also be formed on the gate stack. Optical lithography and dry etching would be performed to define the gate pattern. Implantation of the source/drain region is not intrinsically self-aligned to the gate, and additional masking steps for separate source/drain implantations are needed (Figure 2.2). After gate pattern definition, source extension doping would be carried out with drain side masked. Similarly, drain extension implantation would be carried out with source side masked. Silicon nitride would be deposited using chemical vapor deposition and dry-etched to form spacers. Two more masking and implantation steps would be carried out on the source and drain separately, to form the heavily doped source/drain.

Activation would be carried out after the gate hard mask was removed. Silicidation could be carried out for direct probing on the fabricated devices. Alternatively, passivation layer and metal-1 layer could be deposited and patterned for device protection and electrical measurement.

A more complete process flow is schematically depicted in Figure 2.4, showing the key steps and the device cross-section in each step, corresponding to the steps described above.



Figure 2.4. Key process steps and the corresponding device cross-sections in a typical lateral tunneling FET process flow.
# 2.2 Key Process Challenges for Tunneling FET Fabrication

In order to achieve sub-60 mV/decade *S* in tunneling FETs, certain process requirements need to be considered. These process requirements include formation of abrupt doping profile for the source-to-channel junction and achievement of thin Equivalent Oxide Thickness (EOT) for the gate dielectric. Integration of tunneling-favorable junction is another important requirement. In a tunneling-favorable junction, the material band gap alignment should enhance band-to-band tunneling when a tunneling FET is turned on. By fulfilling the abovementioned requirements, the band-to-band tunneling rate could be enhanced, and the drive current could be increased.

The requirements will be discussed in the following sub-sections, mostly through MEDICI simulations. MEDICI has its own limitations, as will be discussed in Chapter 3. However, it is a readily available tool, and its results are capable of providing useful design guidance. It was used in the preliminary study of tunneling FETs. Base structure simulated is a SOI/SGOI TFET with a body thickness of 40 nm and a gate length of 50 nm.

# 2.2.1 Doping Profile Requirement

Doping profile in the source side affects the electrical characteristics of tunneling FETs. A higher source doping concentration is needed to achieve a higher lateral electric field near the source side of the TFET and to give a lower threshold voltage. The doping junction profile is also an important parameter. Generally speaking, the more abrupt the source-to-channel doping profile, the better the performance of a TFET, in terms of both *S* and on-state current ( $I_{on}$ ), as will be shown in the next page.



Figure 2.5. MEDICI simulation results of a typical TFET device, showing that the source-tochannel junction abruptness affects the device performance in terms of S as well as  $I_{on}$ . Junction abruptness of 1, 2, and 5 nm/decade gives S of ~28 mV/decade, ~27 mV/decade, and ~45 mV/decade, respectively.

MEDICI simulation was used to investigate the effect of junction abruptness on TFET device performance. The  $I_D$ - $V_G$  characteristics of TFET devices with different source doping abruptness is shown in Figure 2.5. The lateral doping abruptness is defined in terms of the change in the horizontal distance needed for the doping concentration to change by ten times. With all other device parameters set to the same values, the source doping abruptness was varied from 5 nm/decade to 2 nm/decade and to 1 nm/decade. We can see that for the device with a more abrupt source-to-channel junction, the threshold voltage is smaller. The threshold voltage was extracted by constant current method, with the constant current set at  $1 \times 10^{-4}$  mA/µm. The subthreshold slope is also steeper for a device with more abrupt source doping. The steeper subthreshold slope gives a smaller *S* value.

The simulation suggests that in order to realize TFET with smaller subthreshold swing, the source-to-channel junction has to be kept as abrupt as possible, preferably below 2 nm/decade. Advanced doping techniques, such as ultra-low energy implantation or cluster ion implantation might need to be utilized in order to form a shallow and abrupt junction. Advanced implantation with dopant diffusion inhibitor might help in reducing dopant diffusion. For example, carbon could inhibit boron diffusion in Si substrate and is a dopant diffusion inhibitor [2.5].

The total thermal budget incurred in all process steps after source implantation needs to be kept low and well controlled, such that dopant diffusion can be minimized. Dopant activation in this work is typically carried out using conventional rapid thermal processing (RTP) which tends to incur a high thermal budget. Such an activation scheme would face the challenges in realizing an abrupt junction. Advanced activation technology such as laser annealing and/or flash anneal would help in realizing a box-like dopant profile [2.6], [2.7]. Epitaxial techniques with *in situ* doping capability would also reduce the thermal budget required, and help achieve a box-like abrupt junction profile.

#### 2.2.2 EOT Requirement

The conventional MOSFET relies on the application of a gate voltage to modulate the barrier height seen by carriers in the source, and to control the carrier flow between source and drain. Similarly, in a TFET, the applied gate voltage controls the tunneling barrier width through modulation of the channel potential, and therefore controls the tunneling current from the source. In order for steep subthreshold swing to be achieved in TFETs, effective gate-to-channel coupling is crucial. The conduction band in the channel is to be pulled down below the valence band in the source in order for band-to-band tunneling to occur. The effectiveness of gate control of the channel potential determines the gate control of the band-to-band tunneling rate.

The gate capacitance plays an important role in the operation of the tunneling FET. A smaller gate dielectric thickness or EOT leads to a higher gate capacitance, and therefore a stronger gate-to-channel potential coupling. However, it has to be noted that gate leakage due to direct tunneling current through the gate dielectric is higher for a thinner gate dielectric. A low off-state leakage current is needed for tunneling FETs for low power consumption, and conventional silicon oxynitride (SiON) gate dielectric with small physical thickness might not be suitable as the gate leakage current may be excessive. The gate leakage current can mask out the low source-to-drain leakage current of the reverse-biased p-i-n diode in a TFET. In order to achieve a thin EOT while keeping the gate leakage current low, a high- $\kappa$  material would be needed. For a given EOT, the physical thickness of a gate dielectric layer will be thicker for a higher  $\kappa$  value, as seen



Figure 2.6. Effect of EOT on tunneling FET performance. It is observed that tunneling FETs with smaller EOT have smaller S and larger on-state current, the S values for the devices with EOT 1 nm, 2 nm, 3 nm is 35mV/decade, 52 mV/decade and 73 mV/decade respectively, estimated over the lowest 6 orders of current change.

from

$$t_{High-\kappa} = \frac{\kappa_{High-\kappa}}{\kappa_{SiO_2}} \times t_{SiO_2},$$
 Equation 2.1

where  $\kappa$  is the dielectric constant and *t* is the physical thickness of a dielectric layer. A larger physical thickness would significantly reduce the gate tunneling leakage current. A high  $\kappa$  dielectric will be used in this work.

Figure 2.6 shows the effect of varying the EOT on the  $I_D$ - $V_G$  characteristics of a tunneling FET. This study was done using MEDICI simulation. It can be observed that a smaller EOT will help in reducing the *S* and increasing the onstate current. EOT below 2 nm would be required in realizing TFET device with reasonable device performance.

# 2.2.3 Material System Requirement

#### 2.2.3.1 Small band gap material

For simplicity, Hurkx's model [2.8] can be used for a qualitative analysis of band-to-band tunneling. Hurkx's model is a modified version of Kane's model [2.9]. In Hurkx's model, the band-to-band generation rate can be expressed by

$$G_{BTBT} = D_{tunnel} \bullet A_{BTBT} \bullet \frac{E^{C_{BTBT}}}{E_g^{1/2}} \bullet \exp\left(-B_{BTBT} \frac{E_g^{3/2}}{E}\right), \qquad \text{Equation 2.2}$$

where  $G_{BTBT}$  is the band-to-band generation rate,  $D_{tunnel}$  is the tunneling factor that takes into account the probability of having a filled state to tunnel from and an empty state to tunnel to.  $A_{BTBT}$ ,  $B_{BTBT}$  and  $C_{BTBT}$  are material related parameters.  $E_g$  is the material band gap.  $D_{tunnel}$  is typically expressed as follows:

$$D_{tunnel} = \frac{1}{1 + \exp\left(\frac{E_{V,1} - E_{Fp,1}}{kT}\right)} - \frac{1}{1 + \exp\left(\frac{E_{C,2} - E_{Fn,2}}{kT}\right)} , \qquad \text{Equation 2.3}$$

where  $E_{V,1}$  and  $E_{Fp,1}$  are the valence band energy and hole Fermi energy on one side of the tunnel barrier, and  $E_{C,2}$  and  $E_{Fn,2}$  are the conduction band energy and electron Fermi energy on the other side of the tunneling barrier.

It can be seen that the material band gap plays an important role in determining the band-to-band generation rate. Utilizing a material with a smaller band gap could greatly increase the band-to-band generation rate and thus increase the on-state current for tunneling FETs. Ge is one of the promising candidates in realizing tunneling FET with a drive current that is high enough to be competitive with conventional CMOS logic devices. According to device simulation, after

optimizing the doping profiles, Ge tunneling FETs with 1.2 mA/ $\mu$ m of on-state current at a supply voltage of 1.2 V could be achieved [2.10].

Other materials like III-V compounds with smaller band gaps would also be advantageous in realizing high performance tunneling FETs at low supply voltage. Such materials include InSb and InN. However, it has to be noted that III-V materials processing is not generally compatible with conventional Si based technology, and there are quite a few process challenges to be surmounted for III-V device fabrication. So far, SiGe and/or Ge based TFET is more promising and more realistic for co-integration with existing CMOS technology. SiGe is selected for device realization in this work.

# 2.2.3.2 Hetero-junction structures

By utilizing materials with smaller band gaps, the on-state current of a tunneling FET could be greatly improved. However, tunneling FET is an ambipolar device, off-state leakage current would also be increased, as band-to-band tunneling could also occur at the drain-to-channel junction at negative gate bias. This ambipolar behavior needs to be suppressed. Besides utilizing materials with smaller band gaps, hetero-junction devices could also be adopted. In the on-state, band-to-band tunneling only occurs at the source-to-channel junction. By adopting a small band gap material as the source material, the device on-state current could be improved. The channel and drain regions are made with materials with larger band gap, so that drain side tunneling leakage could be suppressed. Such concepts have been proposed and analyzed using TCAD tools by Patel *et al.* [2.11] and Toh *et al.* [2.12].



Figure 2.7. Plot of simulated  $I_D$ - $V_G$  characteristics for a single-gate tunneling FET with SiGe source. Increasing the Ge content enhances on-state current and subthreshold swing, while off-state leakage is limited by the drain side, the *S* values for the devices with 40% Ge, 20% Ge, pure Si at source is ~28 mV/decade, ~32 mV/decade and ~38 mV/decade respectively, estimated over the lowest 6 orders of current change.

Figure 2.7 shows the  $I_D$ - $V_G$  characteristics of a simulated SiGe source tunneling FET. It can be seen that increasing the Ge content in the source would improve the on-state current as well as the subthreshold swing. The off-state current is limited by the channel and drain junction. It is kept low and relatively invariant with respect to a change in Ge content in the source. Similar results were reported by Patel *et al.* [2.11] and Toh *et al.* [2.13].

Experimental realization of such a concept was recently reported by S. H. Kim *et al.* [2.14]. By recess etching of the source region followed by a selective growth of polycrystalline Ge in the recess-etched source, hetero-junction TFET was realized, and a record high  $I_{on}/I_{off}$  ratio of up to 10<sup>6</sup> for ultra low voltage (0.5 V) operation was demonstrated. However, the on-state current (reported as a few

 $\mu$ A/ $\mu$ m) is still significantly lower than that of CMOSFETs, whose on-state current is at the level of 1 mA/ $\mu$ m.

#### **2.2.4 Device Body Dimension**

The body thickness of a TFET affects the distribution of electric potential in the channel region of the device. This in turn affects the band-to-band tunneling rate at a given bias. If a thinner Silicon-on-Insulator (SOI) layer was used for device fabrication, a stronger gate-to-channel coupling would be achieved, leading to higher band-to-band tunneling rate. However, when the body thickness is scaled too aggressively, the drive current may be reduced, as the volume of the tunneling region would be effectively reduced. This was shown through simulation for a double-gate TFET [2.15]. Thus, device optimization should be considered carefully for ultra-thin SOI TFETs.

#### 2.2.5 Alignment Issue

In order to fabricate the lateral p-i-n structure, the source and drain regions of a tunneling FET are to be implanted by different dopant types. Individual masking of drain and source side is required, so that the source and drain region can be doped separately. This is typically done by photolithography, with a patterned photoresist covering the source (or drain) region up to the middle of the gate electrode. The drain (or source) region would be exposed separately for implantation. Such a process step imposes stringent requirement for lithography alignment accuracy when the gate length is short. The lithography misalignment needs to be kept smaller than half of the smallest gate length. This requirement can probably be achieved with the state-of-the-art lithography tool, but is very



Figure 2.8. SEM images of devices with the drain side masked. A device with a large gate length  $L_G$ , and properly masked is shown at the top. A device with small  $L_G$  is shown at the bottom. These two devices were subjected to the same amount of lithography misalignment. The misalignment of the drain side masking step leads to a part of the drain region being exposed in the device with short  $L_G$ . This is expected to lead to implantation of source-type dopant into the drain.

challenging for a 248 nm lithography tool used in this work with aggressively trimmed gate lengths.

The best alignment accuracy achievable is around 50 nm. It makes the fabrication of tunneling FET with sub-100 nm gate length very challenging. The non-self-aligned implantation is a critical process for the integration of tunneling FET in an IC with aggressively scaled pitch.

In order to illustrate the severity of lithography misalignment, SEM pictures of two lateral TFET devices before source implantation are shown in Figure 2.8. The device with a larger gate length has a larger tolerance for misalignment. For the device with smaller gate length, the tolerance becomes very tight. With the same amount of lithography misalignment, the p-i-n structure in the device with a larger gate length (top) could be successfully formed, while the p-i-n structure will not be formed in the device with a smaller gate length (bottom) as the drain side is not properly masked for source implantation.

# 2.3 High-ĸ Metal Gate SiGe Tunneling FET

The key process requirements for successful fabrication of a tunneling FET were discussed in the previous sections. A high- $\kappa$  material with a larger dielectric constant would help in realizing tunneling FET devices with good performance. A source/channel material with smaller band gap would help in enhancing band-to-band tunneling rate. By integrating the abovementioned factors, high- $\kappa$  metal gate SiGe tunneling FET devices were fabricated, as to be described next.

# 2.3.1 Experimental Realization of SiGe Tunneling FET

In this experiment, a fabrication process employing 8-inch bulk Si (001) substrates was adopted. Local Oxidation of Silicon (LOCOS) process was used to isolate the active regions. 250 nm of wet oxide was grown with the active regions masked by a silicon nitride hard mask. The silicon nitride hard mask was then removed by wet etch.

Standard pre-gate cleaning process was carried out, including an SPM (H<sub>2</sub>SO<sub>4</sub>:H<sub>2</sub>O<sub>2</sub> mixture solution) cleaning, an SC-1 (solution of Ammonium hydroxide and hydrogen peroxide) cleaning and a hydrofluoric acid (HF) dip. The SPM cleaning removes the organic residues. The SC-1 cleaning removes metallic contamination. The HF dip removes native oxide.

After the pre-gate cleaning process, about 20 nm of  $Si_{0.75}Ge_{0.25}$  was selectively grown on the Si active area using an ultra-high vacuum chemical vapor deposition system. Since band-to-band tunneling only occurs in the surface layer near the gate, the SiGe layer would act as the active layer. Around 3 nm of Si cap

was grown on top of the SiGe layer, in order to improve the channel to dielectric interface quality.

Another HF dip was carried out before the gate dielectric was deposited. 5 nm of  $HfO_2$  high- $\kappa$  dielectric was deposited by atomic layer deposition (ALD) technique. 80 nm of TaN was deposited by sputtering to form the gate.

Gate stack was then patterned and etched, source/drain implantation and activation was carried out as described in the previous sections. The fabricated devices were subjected to direct probing for electrical characterization. The device states in some of the critical steps are shown in Figure 2.9.



Figure 2.9. The SiGe tunneling FET active area (left) and a device with drain side masked for source side implantation (right).

#### 2.3.2 Results and Discussion



Figure 2.10. TEM image of a SiGe tunneling FET, featuring 5 nm of  $HfO_2$  as the gate dielectric, and  $Si_{0.75}Ge_{0.25}$  as the active layer.

Transmission Electron Microscopy (TEM) was used to study the fabricated devices. TEM image of a device is as shown in Figure 2.10. Crystalline SiGe was observed in the active region with a thickness of about 20 nm. Energy-dispersive X-ray (EDX) spectroscopy was used to analyze the Ge concentration. Around 25% to 30% of Ge concentration was found in the SiGe layer. The gate dielectric thickness was measured to be around 5 nm.

By using  $Si_{0.75}Ge_{0.25}$  as the active material instead of pure Si, the energy band gap can be reduced from 1.14 eV down to around 0.95 eV, as seen from Figure 2.11. Due to the fact that band-to-band tunneling is directly affected by the band gap of the material system, using a material with a smaller band gap, e.g. SiGe, is expected to enhance the tunneling rate and the drive current of the TFET as compared with a Si TFET.



Figure 2.11. Composition dependence of energy band gap in Ge-Si alloys at room temperature [2.17].

Dielectric quality of the ALD deposited HfO<sub>2</sub> (done in SNDL) was examined through a capacitor fabrication experiment. A two-mask capacitor run was used. The active area pattern was defined by photolithography on test wafers with 400 nm of field oxide. Buffered oxide etch (BOE) was used to open the active area. After standard pre-gate cleaning, 5 nm of HfO<sub>2</sub> dielectric was grown by ALD process. 100 nm of TaN metal gate was deposited on top of the dielectric to form the metal gate. The metal gate layer was deposited by sputtering. Gate lithography and dry etch were used to define the capacitor gate contact.

High frequency capacitance-voltage (*C*-*V*) measurement was carried out on the fabricated capacitors. A *C*-*V* simulator [2.16] which accounts for quantum mechanical effects was used to fit the measurement data. EOT and the flat-band voltage were extracted through fitting the simulated curve with the experimental curve. For capacitors annealed at different temperatures, 1.2 nm to 1.6 nm of



Figure 2.12. Measured *C-V* characteristics of fabricated high- $\kappa$  metal gate capacitors. Fitting of a simulated *C-V* curve which accounted for quantum mechanical effects to an experimental curve was carried out to obtain the Equivalent Oxide thickness (EOT) and the flatband voltage. The plot on the top is for a capacitor which underwent a 550 °C 30 s post-deposition anneal (PDA). It is fitted with a flat-band voltage ( $V_{fb}$ ) of -0.4 V and an EOT of 1.2 nm. The bottom plot is for a capacitor which underwent a 600 °C 30 s PDA. It is fitted using  $V_{fb}$  of -0.4 V and an EOT of 1.4 nm.

EOT can be obtained. The flat-band voltage was fitted to be around 0.4 V. The gate leakage current was found to be low too. A typical fitting result is shown in

Figure 2.12. It is expected that the dielectric used in the SiGe tunneling FET would have the same quality as that of the capacitors.

Tunneling FETs with sub-1.5 nm of EOT were fabricated. The fabricated SiGe tunneling FETs with different gate lengths were electrically characterized with the p-type region (drain) grounded and a positive bias applied to the n-type region (source), i.e. p-type TFET operation. The  $I_D$ - $V_G$  curves are plotted in Figure 2.13.

However, the device performance is not as good as expected. The best S values are found to be only around 300 mV/decade. The on-state current are found to be worse than that shown in the simulated results. This might be attributed to the poor source doping profile.



Figure 2.13.  $I_D$ - $V_G$  characteristics of the fabricated SiGe high- $\kappa$  metal gate tunneling FETs. Devices with different gate length are shown. The tunneling current is not dependent on the gate length, as expected.

Figure 2.14 shows the simulated doping profile in the source along the vertical direction. Implantation and dopant activations conditions are similar to those used in the actual device fabrication. It can be seen that the doping abruptness is very poor, with a value of up to around 20 nm/decade. The lateral doping profile is expected to be about 20 nm/decade. This might be the reason why the *S* value for the fabricated device is not good and the on-state current is low. The gentle doping profile limits the lateral electric field built up across the tunneling junction. The tunneling width would be increased as a result. Simulation results presented in the previous section showed that a sub-2 nm/decade doping abruptness is required in order to realize tunneling FET with good performance. Advanced techniques that can help in realization of more abrupt junction might help in improving the device performance.

It is interesting to note that devices with different gate lengths have similar  $I_D$ - $V_G$  characteristics. The off-state leakage current and on-state drive current are found to be nearly independent of the gate length. The off-state leakage current  $I_{off}$  was found to be low and no trend of significant increase in  $I_{off}$  was observed while the gate length was further reduced, as shown in Figure 2.15. This indicates that the tunneling junction is the limiting factor for current conduction. Effect of channel resistance caused by gate-length difference is not significant. No significant device performance difference was found for devices with gate lengths from 5 µm down to 180 nm. Scaling of tunneling FET devices has little effect on their off-state current. This is advantageous as compared to the conventional MOSFET, whose off-state current increases with device scaling due to increased short channel effects. Performance of the fabricated SiGe tunneling FET was

limited by the source doping profile. The experimental efforts in forming an abrupt doping profile in tunneling FETs will be discussed in the following section.



Figure 2.14. Simulated source doping profile along the vertical direction.



Figure 2.15. The off-state leakage of the fabricated SiGe tunneling FET is low and independent of the gate length. No trend of significant off-state current increase was observed while the gate length was reduced.

# 2.4 Dopant Segregation (DS) Technique and Its Potential Application in Tunneling FETs

After analyzing the key process challenges and the typical process flow in the fabrication of tunneling FETs, we realize the limitations of existing process capabilities and considered a new way to surmount the challenges. It is known that the dopant segregation technique is useful for forming very abrupt doping profile with high peak doping concentration. We therefore conceived that the dopant segregation technique could be helpful for forming an abrupt source dopant profile in the TFET, and would be useful for realizing TFETs with good performance.

R. L. Thornton is one of the earliest researchers working on the Dopant Segregation technique (DS) for contact formation [2.18]. The technique involves silicidation of an implanted region. The silicide layer is chosen, such that it is thick enough to fully consume the doped or implanted Si layer. Depending on the material system, the dopant would either be redistributed in the silicide layer, or be piled-up at the silicide-substrate interface. When there is dopant segregation, the profile of the segregated dopants can be more tightly distributed than that of the as-implanted dopant profile, and the peak concentration of the segregated dopants can be higher than that of the as-implanted dopant profile. Dopant segregation can therefore help to achieve higher doping concentration, as well as a more abrupt dopant profile. This technique could be helpful in realizing tunneling FETs with good performance. This new idea was explored as follows. By applying dopant segregation technique, a segregated and heavily doped source region would be formed at the silicide-Si interface. A higher doping concentration in the source is expected. A more abrupt junction between this segregated source and channel region could help in enhancing band-to-band tunneling rate, due to the higher built-in electric field. An ohmic contact comprising a tunneling junction through a Schottky barrier in series with a bandto-band tunneling junction is depicted. The dopant-segregated ohmic contact has a low contact resistance and is not expected to be a significant limiting factor even for high drive currents observed in high performance MOSFETs. Band-to-band tunneling FET characteristics is expected. This idea is depicted in Figure 2.16.

Experimental investigation on dopant pile-up at the silicide-semiconductor interface was reported in [2.19]. Figure 2.17 is taken from [2.19], where both boron and arsenic were reported to pile up due to NiSi induced segregation. The segregated dopant profile was found to be of non-negligible thickness, this could form the heavily doped source region if DS is applied in fabrication of tunneling FETs. Similar observation was made by Kinoshita *et al.* and Wong *et al.* [2.20]-[2.21] where MOSFETs was fabricated using different material systems.



Figure 2.16. Idea for application of dopant segregation in the fabrication of tunneling FETs. Shown on the top is a schematic of a DSS tunneling FET. The tunneling junction is zoomed-in and shown in the middle, with the corresponding band diagram along the dashed line shown at the bottom. Current conduction is controlled by the Schottky tunneling junction in series with the band-to-band tunneling junction. The metal-to-semiconductor contact would behave as an ohmic contact attributed to the elevated doping concentration in the segregated layer. Due to the low band-to-band tunneling rate, it is expected to be the current limiting factor. Tunneling FET behavior is expected.



Figure 2.17. Boron (B) profiles of DS junctions using  $CoSi_2$  (left) and NiSi (right), taken by backside SIMS [2.19].

# 2.4.1 Experimental Procedure

This work, done in year 2009, was among the first experimental efforts on integrating the DS technique in tunneling FET fabrication. Ni silicidation on

doped Si substrate was performed. 8-inch Silicon-on-Insulator (SOI) wafers were used as the starting substrates. The initial SOI body thickness is 70 nm, on top of 145 nm of buried oxide layer. In order to make the tunneling junction near the gate, effective silicidation encroachments towards under the gate is required. SOI thinning steps were carried out to enable effective silicide encroachment. The SOI thinning steps was done by thermal oxidation and oxide wet strip. SOI thinning steps were repeated, and a final Si layer thickness of around 20 nm was achieved.

Mesa was patterned and etched to form the active area of the devices. After pre-gate cleaning, thermal oxidation at 850 °C was used to form 30 Å of gate dielectric. Dry oxidation was used to achieve good gate dielectric quality.



Figure 2.18. An SEM image of the ultra-thin body tunneling FET with drain side masked.

Polycrystalline Si was deposited on the gate dielectric to form the gate electrode. The standard tunneling FET process flow was followed. Since the Si body is ultrathin, the deep source/drain implantation steps were skipped. P<sup>+</sup> region was doped with  $BF_2^+$  at a dose of  $1 \times 10^{15}$ /cm<sup>-2</sup> and an implant energy of 7 keV. N<sup>+</sup> region was doped with As<sup>+</sup> at a dose of  $5 \times 10^{14}$  /cm<sup>-2</sup> and an implant energy of 6 keV. An SEM image of the device before source implantation is shown in Figure 2.18. Dopant activation was carried out with a *Jipelec Jetstar* Rapid Thermal Process (RTP) system, at 950 °C for 30 seconds. The silicon nitride spacers were measured to be around 20 nm thick.

After dopant activation, Ni was deposited on the device, and silicidation process was carried out with a RTP process at 450 °C for 30 seconds. SPM cleaning was followed to remove any un-reacted Ni from the device.

#### 2.4.2 Results and Discussion

Before the DS experiments were carried out, devices silicided with a thin layer of Ni was fabricated and characterized. They were proven to work as tunneling FETs. 5 nm of Ni was deposited on the devices using an electron beam physical vapor deposition (E-beam evaporator). The SOI body thickness was measured to be 20 nm, and the spacers were around 20 nm thick. It is believed that the Si body is only partially consumed and no dopant segregation is expected at the tunneling junction near the gate edge.

Figure 2.19 shows the typical *I-V* characteristics of a fabricated tunneling FET. The device drain side ( $n^+$  region) is biased at 2.5 V, the source side ( $p^+$  region) is grounded. Gate voltage is swept from -4 V to 4 V in steps of 50 mV.



Figure 2.19. A typical  $I_D$ - $V_G$  characteristics of a fabricated device silicided with 5 nm of Ni. The device gate length is around 1  $\mu$ m, device width is around 0.8  $\mu$ m, device drain side is biased at 2.5 V.

Ambipolar conduction behavior was observed. The gate leakage is negligible as compared with the drain current. The fabrication process and the measured *I-V* characteristics indicated that this is a working tunneling FET, and the gate dielectric is of good quality. Note that the p-channel TFET operation (where  $V_{gs} < 0$ ) is better than that of n-channel TFET operation (where  $V_{gs} > 0$ ). This can be attributed to the fact that As diffusion in Si is much lower than that of B in Si, thus a more abrupt As profile was expected. The discussion will focus on the p-channel TFET operation.

The device performance is not as good as theoretically predicted. The best *S* value is found to be a few hundred mV/decade. The on-state current is much lower compared with state-of-the-art CMOSFET results as well.

As the control device fabricated is a TFET, dopant segregation was applied to the source and drain regions to investigate its effect on the device performance. 15 nm of Ni was deposited on the device and silicidation was carried out by annealing at 450 °C for 30 seconds. Standard SPM cleaning was used to remove un-reacted Ni from the spacers. It is believed that the Si source and drain would be fully consumed by the silicidation process. Significant amount of silicide encroachment towards under the gate is expected.

The measured  $I_D$ - $V_G$  characteristics are as shown in Figure 2.20. The gate leakage remained as low, as compared to the drain current. Two distinct subthreshold slopes were observed for the p-channel TFET operation. An *S* value of around 70 to 80 mV/decade was observed at low gate voltages. An *S* value of a few hundreds of mV/decade was observed for negative gate bias.



Figure 2.20.  $I_D$ - $V_G$  characteristics of a TFET device silicided by 15 nm of Ni. The device gate length is around 2  $\mu$ m, device width is 0.8  $\mu$ m, device drain side is biased at 2.5 V.

This characteristic corresponds to a Schottky barrier (SB) FET as discussed by J. Knoch *et al.* [2.22]. This could happen if the segregated dopant is distributed within a very thin layer, and being fully depleted. In such a case, there would not be a tunneling source available for band-to-band tunneling to occur. Direct Schottky tunneling across the Schottky barrier would be the only current conduction mechanism.

Figure 2.21 is taken from [2.23]. Energy Dispersive X-ray (EDX) analysis was used to analyzed the Co-induced As segregation as the silicide interface. The segregated dopant is as indicated by the bright band at the interface, indicating a very small width of the segregated dopant profile.



Figure 2.21. Atomic resolution Z-contrast images in cross section of an As-doped device, silicided using Cobalt [2.23].

Similar effect might have happened within the Ni silicided tunneling FET device. The deposited Ni layer has consumed the entire doped Si layer. The segregated dopant layer is too thin to form the tunneling source, resulting in a SB FET instead of a tunneling FET.

In order to make the device perform as a tunneling FET, very stringent silicide process control is required. The silicidation process should be controlled such that not all the doped Si is consumed, yet a highly doped peak could be formed during the dopant segregation. The boosted peak doping concentration could also create steeper doping profile, thus help in enhancing band-to-band tunneling.

However, such a process is very challenging as far as process control is concerned, and large variability in device performance is expected if process uniformity is not well-controlled. This process might still attract some research interest due to its potential in boosting TFET device performance.

# 2.5 Summary

In this chapter, challenges in the fabrication of tunneling FET devices were discussed. In order to fabricate tunneling FET devices with small *S* value and high on-state current, the process requirements are very stringent. Thin EOT with low gate leakage current is required to increase gate-to-channel coupling. High- $\kappa$  dielectric is useful for suppressing gate leakage current. Materials with smaller band gap would enhance the band-to-band tunneling rate. This indicates that Ge/SiGe material could be useful. Very abrupt tunneling junction doping profile is required, which makes advanced doping technique a must.

Experimental efforts in fabrication of tunneling FET device were carried out. SiGe tunneling FETs with metal gate and high- $\kappa$  gate dielectric were fabricated. The fabricated devices were found to behave as tunneling FETs, while the device performance is not as good as expected. This is attributed to specific process limitations. High source doping concentration with abrupt doping profile was not achieved. Dopant segregation technique was used in tunneling FET fabrication. Advanced process development and tight process control are required in order to improve the fabricated device performance.

In the next chapter, simulation study on tunneling FET device will be discussed. A novel non-local band-to-band tunneling algorithm was developed and applied in tunneling FET simulation study.

# **Reference:**

- [2.1] A. S. Verhulst, *et al.*, "Complementary silicon-based heterostructure tunnel-FETs with high tunnel rates," *IEEE Electron Device Letters*, vol. 29, pp. 1398-401, 2008.
- [2.2] T. Nirschl, P.-F. Wang, W. Hansch and D. Schmitt-Landsiedel, "The tunneling field effect transistor (TFET): The temperature dependence, the simulation model, and its application," Vancouver, BC, Canada, 2004, pp. III713-III716.
- [2.3] A. S. Verhulst, et al., "Complementary silicon-based heterostructure tunnel-FETs with high tunnel rates," *IEEE Electron Device Letters*, vol. 29, pp. 1398-401, 2008.
- [2.4] A. S. Verhulst, W.G. Vandenberghe, K. Maex and G. Groeseneken, "Boosting the on-current of a n -channel nanowire tunnel field-effect transistor by source material optimization," *Journal of Applied Physics*, vol. 104, 2008.
- [2.5] Rucker, H., B. Heinemann, *et al.* (1998). "Suppressed diffusion of boron and carbon in carbon-rich silicon." *Applied Physics Letters*, 73, 1682-1684.
- [2.6] D. X. M. Tan, et al., "Vacancy engineering by optimized laser irradiation in boron-implanted, preamorphized silicon substrate," Applied Physics Letters, vol. 92, 2008.
- [2.7] J. Gelpey, et al., "Ultra-shallow junction formation using flash annealing and advanced doping techniques," in 2008 8th International workshop on Junction Technology (IWJT '08), 15-16 May 2008, Piscataway, NJ, USA, 2008, pp. 84-8.
- [2.8] G. A. M. Hurkx, D. B. M. Klaassen and M.P.G. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Transactions on Electron Devices*, vol. 39, pp. 331-8, 1992.
- [2.9] E. O. Kane, "Theory of Tunneling," *Journal of Applied Physics*, vol. 32, p. 83, 1961.
- [2.10] E.-H. Toh, G. H. Wang G. Samudra and Y-C. Yeo, "Device physics and design of germanium tunneling field-effect transistor with source and drain engineering for low power and high performance applications," *Journal of Applied Physics*, vol. 103, pp. 104504-1, 2008.
- [2.11] N. Patel, A. Ramesha and S. Mhapatra, "Drive current boosting of n-type tunnel FET with strained SiGe layer at source," *Microelectronics Journal*, vol. 39, pp. 1671-7, 2008.
- [2.12] E.-H. Toh, G. H. Wang G. Samudra and Y-C. Yeo, "Device physics and guiding principles for the design of double-gate tunneling field effect transistor with silicon-germanium source heterojunction," *Applied Physics Letters*, vol. 91, pp. 243505-1, 2007.
- [2.13] E.-H. Toh, G. H. Wang G. Samudra and Y-C. Yeo, "Device design and scalability of a double-gate tunneling field-effect transistor with silicon germanium source," *Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers*, vol. 47, pp. 2593-2597, 2008.
- [2.14] S. H. Kim, Kam He, C. Hu and T-J. King-Liu, "Germanium-Source Tunnel Field Effect Transistors with Record High I<sub>on</sub>/I<sub>off</sub>," in *Symposium* on VLSI Technology Digest of Technical Papers, 2009, p. 178.
- [2.15] E.-H. Toh, G. H. Wang G. Samudra and Y-C. Yeo, "Device physics and design of double-gate tunneling field-effect transistor by silicon film

thickness optimization," Applied Physics Letters, vol. 90, pp. 263507-1, 2007.

- [2.16] "UC Berkeley Device Group, This Software calculates the capacitancevoltage characteristics of an MOS capacitor considering the electron/hole distribution in both iversion and accumulation. These distributions are calculated by solving Schroedinger's and Poisson's equations selfconsistently with the Fermi-Direc distribution using the phsycal parameters for (100) Si. ."
- [2.17] R. Braunstein, "Intrinsic Optical Absorption in Germanium-Silicon Alloys," *Physical Review*, vol. 109, p. 695, 1958.
- [2.18] R. L. Thornton, "Schottky-barrier elevation by ion implantation and implant segregation," *Electronics Letters*, vol. 17, pp. 485-6, 1981.
- [2.19] J. Knoch, M. Zhang, Q.T. Zhao, U. Breuer and S. Mantl, "Impact of dopant segregation on fully depleted Schottky-barrier SOI-MOSFETs," *Solid-State Electronics*, vol. 50, pp. 594-600, 2006.
- [2.20] A. Kinoshita, Y. Tsuchiya, A. Yagishita, K. Uchida and J. Koga, "Solution for high-performance Schottky-source/drain MOSFETs: Schottky barrier height engineering with dopant segregation technique," Honolulu, HI, United states, 2004, pp. 168-169.
- [2.21] H.-S. Wong, L. Chan, G. Samudra and Y-C. Yeo, "Low Schottky barrier height for silicides on n-type Si (100) by interfacial selenium segregation during silicidation," *Applied Physics Letters*, vol. 93, pp. 072103-1, 2008.
- [2.22] J. Knoch, M. Zhang, S.Feste and S. Mantl, "Dopant segregation in SOI Schottky-barrier MOSFETs," *Microelectronic Engineering*, vol. 84, pp. 2563-71, 2007.
- [2.23] T. Topuria, N. D. Browning and Z. Ma, "Characterization of ultrathin dopant segregation layers in nanoscale metal-oxide-semiconductor field effect transistors using scanning transmission electron microscopy," *Applied Physics Letters*, vol. 83, pp. 4432-4434, 2003.

# Chapter 3

# A Non-Local Band-to-Band Tunneling Algorithm and its Application in Tunneling FET Variability Study

Device simulation is a fast and convenient way to study the physics of new semiconductor devices. A well established simulation model can simulate a device with reasonably good accuracy. By utilizing such device simulation models, device characteristics can be studied without performing costly device fabrication, making it very time efficient as well. Parameter isolation could be achieved, i.e. the effect of varying each design or physical parameter on the device performance can be analyzed separately, with the effect due to variation of all other parameters excluded. Such an advantage could greatly help in device understanding and modeling. Tunneling Field Effect Transistor (tunneling FET or TFET) was extensively studied by simulation, due to the fact that numerical simulation provided a fast and effective way to study tunneling FET performance characteristics.

# **3.1 Development of a New Non-local Algorithm for Band-to-Band Tunneling**

In order to study the tunneling FET device accurately, an accurate band-toband tunneling model is required, as it is the dominant and core physical mechanism that captures the device physics of the tunneling FET. However, it has to be noted that the band-to-band tunneling in tunneling FETs is at least two-dimensional (2D). One-dimensional (1D) tunneling is well-established and much easier to model [3.1], [3.2]. However, 2D tunneling is not easy to model accurately. The implementation of the existing models in a physical and accurate manner is not an easy task either. There are commercially available simulation algorithms. However, when problems are encountered in the use of commercial simulation tools, it is often not easy to understand and debug because the algorithms and code used are not accessible. We cannot evaluate and solve the issues faced in using commercial simulators to model the tunneling FET. The fact that the actual algorithm and code used in commercial simulators are not usually disclosed affects the usefulness of applying the simulators to tunneling FET simulation.

Band-to-band tunneling results in the generation of electron hole pairs in regions of high electric field when the band bending is large enough and significant tunneling occurs. E. O. Kane derived a mathematical modeling for band-to-band tunneling [3.1], [3.2]. This model was used in many device simulators. Kane did a detailed calculation for Zener tunneling in a constant field, by adapting a simple two-band model. An inter-band matrix element was calculated with the stationary-phase method [3.1]. Based on his calculation, a famous Kane's model for band-to-band tunneling was developed and widely adopted in recent device simulators like MEDICI. The Kane's model is also used for tunneling FET device simulation by quite a number of researchers [3.3] - [3.10].

Kane's model is defined as follows [3.1], derived based on band-to-band tunneling in InSb,

$$n = \frac{E^2 m_r^{1/2}}{18\pi\hbar^2 E_g^{1/2}} \exp\left(\frac{-\pi m_r^{1/2} E_g^{2/3}}{2\hbar E}\right),$$
 Equation 3.1

where *n* stands for the generation of electrons in units of per second per cm<sup>3</sup>, *E* stands for the electric field,  $m_r$  is a reduced tunneling mass,  $E_G$  is the material band gap, and  $\hbar$  is reduced Planck constant.

The model implemented in MEDICI [3.11] is as follows,

$$G_{BTBT} = A_{BTBT} \frac{E^{C_{BTBT}}}{E_g^{1/2}} \cdot \exp\left(-B_{BTBT} \frac{E_g^{3/2}}{E}\right), \qquad \text{Equation 3.2}$$

where  $G_{BTBT}$  stands for the band-to-band generation rate, which directly affects the band-to-band tunneling current for a given device structure and material system.  $E_g$  is the energy bandgap of the semiconductor material. By comparing the two equations above, it can be seen that  $A_{BTBT}$ ,  $B_{BTBT}$  and  $C_{BTBT}$  are parameters related to the material properties under investigation; these parameters are functions of the relative tunneling mass (or reduced tunneling mass, as defined by Kane), as well as the bandgap of the material under investigation. Kane's model was derived based on the simple twoband theory for direct band gap material like InSb. Constant field is assumed in the tunneling region.

Band-to-band tunneling is a non-local phenomenon. The availability of electrons in the valence band on the source side and empty states in the conduction band on the destination side will affect the band-to-band tunneling rate. Based on Kane's model, Hurkx proposed a modified version of Kane's model [3.12],

$$G_{BTBT} = D_{tunnel} \bullet A_{BTBT} \frac{E^{C_{BTBT}}}{E_g^{1/2}} \bullet \exp\left(-B_{BTBT} \frac{E_g^{3/2}}{E}\right), \qquad \text{Equation 3.3}$$

where an additional  $D_{tunnel}$  factor is considered in the equation. The factor is defined as follows [3.11],

$$D_{tunnel} = \frac{1}{1 + \exp\left(\frac{E_{V,1} - E_{Fp,1}}{kT}\right)} - \frac{1}{1 + \exp\left(\frac{E_{C,2} - E_{Fn,2}}{kT}\right)},$$
 Equation 3.4

where  $E_{V,I}$  and  $E_{Fp,I}$  are the valence band energy and the hole Fermi level on one side of the tunnel barrier respectively, while  $E_{C,2}$  and  $E_{Fn,2}$  are the conduction band energy and electron Fermi level on the other side of the tunnel barrier, respectively. By including this parameter, evaluation of band energy levels on both sides of the tunnel barrier was included in the band-to-band tunneling calculation. Hurkx's model takes into account the probability of having a filled state (electron) to tunnel from and an empty state to tunnel to. This parameter allows for both Zener tunneling and Esaki tunneling, and gives a more accurate evaluation of band-to-band tunneling as compared to the original Kane's model.

It has to be noted that the abovementioned formulation of band-to-band tunneling is typically done on a 1D structure, and parameter fitting (mainly the tunneling mass) is also carried out using a 1D structure only. A realistic tunneling FET structure is three-dimensional (3D), but can be typically simplified to 2D in a device simulation. The band-to-band tunneling direction is not fixed and varies with terminal voltages. Changes in device terminal voltages would change the potential or electric field distribution, leading to changes in the tunneling directions. The abovementioned device physics imply that simulating tunneling FETs using 1D based
tunneling models where the tunneling direction is fixed will have a limited accuracy. To apply such models in a 2D device simulation, various algorithms are used, each of which has its own advantages as well as intrinsic limitations. Simulation of tunneling by adopting these models might not cause much problem in simulation of conventional MOSFETs, as band-to-band tunneling only contributes to a part of the leakage current. However, this issue could be significant in tunneling FETs, as band-to-band tunneling is the main current conduction mechanism. A novel, robust, and non-local algorithm for band-to-band tunneling simulation in multiple directions is desired for reasonably accurate tunneling FET simulations.

Our team developed a novel algorithm for 2D simulation of band-to-band tunneling, and demonstrated its applications in tunneling FET simulation [3.13]. This algorithm is based on an open source device simulator named General-purpose Semiconductor Simulator (GSS) [3.13]. There is a commercially supported version by *Congenda* [3.15]. The implementation of this algorithm and its application in tunneling FET simulation will be discussed in following sections.

As the band-to-band tunneling is a non-local process, it is important to note that the calculation of band-to-band tunneling rate strongly depends on a tunneling distance. This tunneling distance is defined by the starting and ending points of a band-to-band tunneling path, connecting the valence band on one side of the barrier to the conduction band on the other side of the barrier. Carrier concentration on both sides of the barrier affects the tunneling current as well. Based on these considerations, our team developed a tunnel path based algorithm, using Wentzel-Kramers-Brillouin (WKB) approximation [3.13].

By applying the derivation of Huang *et al.* [3.16] for a band-to-band tunneling problem, the most probable tunnel path of an electron can be easily obtained by using Newton mechanics in an inverted potential and energy domain, assuming that electrons strike on the barrier in a normal direction.

The algorithm implementation is carried out in a 2D simulation domain. A device structure is discretized in spatial domain through the construction of a 2D mesh. Physical parameters for the simulation, including device dimensions, doping levels, Equivalent Oxide Thickness (EOT), etc. are included in the input file. This is similar to the approach taken by other TCAD device simulators.

Figure 3.1 shows a typical mesh structure for tunneling FET simulation. Typically, very dense mesh is required in the region where tunneling occurs, in order to simulate the quantum tunneling effects with reasonable accuracy. This is due to the fact that band-to-band tunneling occurs across a very small physical distance. In the simulation, electrical properties were calculated and stored in each of the nodes in the spatial domain.

Besides the discretization in spatial scale as shown in Figure 3.1, energy scale discretization is also carried out. At the tunneling junction, for a certain energy



Figure 3.1. A typical mesh for a gated p-i-n structure for tunneling FET simulation.

level, conduction and valence band front ( $E_C$  front and  $E_V$  front) are searched and constructed on the mesh. The conduction band front is discretized into segments or strips. The algorithm takes the center of each segment as the starting point of a tunneling path. Based on the formulation of Huang *et al.* [3.16], the tunneling path was searched and stored. An imaginary particle is put at a point on the  $E_C$  front. This particle within the "inverted" tunnel barrier moves towards the valence band front at the source region. The particle hits the valence band front and marks the starting position of this tunneling path. Figure 3.2 illustrates the tunneling paths between the  $E_V$  front and  $E_C$  front at a particular energy level, with the blue dashed line starting from the valence band on the source side. Conduction band edge and valence band edge at each different energy level are searched and paired up. Tunneling paths are thus found between each pair of conduction and valence band edges, as shown in Figure 3.3.



Figure 3.2. At a particular energy level, tunneling paths between  $E_C$  front and  $E_V$  front are found on the mesh. Each tunneling path accounts for the tunneling current in a strip defined between  $E_C$  front and  $E_V$  front.



Figure 3.3. Band Diagram of a tunneling junction visualized in 3D. Tunneling paths at a certain energy level are indicated by black arrows.

Band-to-band tunneling rate is calculated during the search of tunneling paths, by assuming that the tunneling current within each strip is independent of each other. The tunneling probability associated with each strip is evaluated numerically through WKB approximation,

$$T_{tun} = \exp\left\{-2\int_{C}^{V} \frac{\sqrt{2m_{r}\left(V - E_{bar}\right)}}{\hbar} dr\right\},$$
 Equation 3.5

where  $T_{tun}$  stands for the tunneling probability, and the integration is taken along the tunneling path from  $E_C$  front to  $E_V$  front.  $m_r$  stands for the effective tunneling mass. V and  $E_{bar}$  stand for the tunneling barrier energy and electron energy, respectively.

During the implementation of the integration, the tunneling path in each of the mesh cell is assumed to be a straight line segment, since the direction of the tunneling path does not change much in each mesh cell.

Based on the projection of the Voronoi cells at the starting and ending point of a tunneling path, the tunneling current could be evaluated through the following formula [3.17].

$$dG_{BTBT} = \frac{qm_{r,dos}kT}{2\hbar^3} \frac{n_c p_v - n_i^2}{\left(n_c + n_i\right)\left(p_v + p_i\right)} W \bullet T_{tun} \bullet dE, \qquad \text{Equation 3.6}$$

where  $n_c$  is the electron concentration at the node right behind  $E_c$  front, and  $p_v$  is the hole concentration at the node behind the  $E_V$  front. Electrons/holes generated by band-to-band tunneling are added to the electron/hole continuity equations at the two nodes. The generation is calculated recursively, until convergence is achieved. Finite volume discretization was applied; Newton's method is adopted to solve the nonlinear equations.

This band-to-band tunneling algorithm implemented with GSS is proven to be useful and robust in solving 2D band-to-band tunneling problem [3.13]. Its applications in simulation of a 2D tunneling diode as well as tunneling FETs will be demonstrated next.

The algorithm was first tested on a 2D tunneling diode. Two different mesh structures were constructed; one with intentionally densified mesh at the tunnel junction in order to simulate the band-to-band tunneling behavior more accurately, the other mesh was generated by the GSS simulator itself, only being re-meshed automatically according to potential and doping profile.

The non-local band-to-band tunneling algorithm works well with simple mesh structure automatically generated by the simulator itself. Figure 3.4 depicts the two mesh structures mentioned in the previous paragraph. In order to capture the band-to-band tunneling effectively, mesh B was intentionally densified. It is much more computationally expensive than that of the auto-generated Mesh A. The calculated *I*-V characteristics are also shown in Fig. 3.4, which coincide with each other. It can be seen that the non-local BTBT algorithm is not mesh dependent.

One of the most advantageous features of the non-local band-to-band tunneling model is that the algorithm is oriented towards the search for possible tunneling paths. By extracting and plotting the tunneling paths in a simulation mesh, information like tunneling direction and tunneling barrier width/distance can be easily visualized. Providing such information is important for band-to-band tunneling simulation. This feature makes it easier to visualize the band-to-band tunneling. It could help in understanding the operation of tunneling FET devices.



Figure 3.4. (a) A mesh for a  $p^+$ -n 2D tunneling diode, generated by GSS simulator. (b) A manually created mesh for the same diode which is denser than that in (a). (c) *I-V* characteristics obtained by simulation the two diodes meshed in (a) and (b), using GSS non-local BTBT algorithm [3.13].

The simulator is shown to be useful for tunneling FET device simulation [2.13]. This algorithm is proven to be versatile in tunneling FET simulation. Based on this algorithm, tunneling FET variation study was carried out, as will be discussed in the following section.

# **3.2** Application of the Algorithm in Tunneling FET Variability Study

Development of the novel non-local band-to-band tunneling algorithm makes the simulation of tunneling field effect transistor more physical and accurate.

Process induced variation is a major limitation factor to miniaturized MOSFETs [3.18]. It is widely believed that parameter variations would drastically limit the prospect of future nanoscale MOSFETs. Emerging devices might also suffer from device to device performance variation. Tunneling FET is considered a novel device concept which might find its application in post-CMOS era. A variability study of the tunneling FET is important in understanding its prospective adoption in mass production.

Tunneling FET performance variation is studied through TCAD simulation. Germanium is selected as the material used for the tunneling FET device, as band-toband tunneling current is expected to be higher in Ge based tunneling FETs than in Si based tunneling FETs. Ge possesses a lower band gap, down to 0.661 eV, as compared to Si, whose band gap is 1.12 eV. Significant improvement is expected for Ge tunneling FETs, as compared to Si tunneling FET devices.

# 3.2.1 Calibration of Tunneling Mass

In order to model band-to-band tunneling in Ge tunneling FET, the tunneling mass in Ge  $(m_r)$  is the most important parameter to be determined. The tunneling probability is directly affected by this parameter based on the WKB approximation calculation, as shown in Equation 3.5.

There are two ways to obtain the tunneling mass to be used in the simulation. It can be either done by fitting of experimental results or done through theoretical derivation.

In this study, the method of experimental data fitting was adopted. The theoretical derived values were used to check whether there is any discrepancy from the fitted data. The accuracy of the tunneling algorithm can be further confirmed if there is no significant discrepancy between the fitted tunneling mass and theoretically derived data. To obtain more accurate fitting, simple tunneling diode structure was chosen. The simple structure would help focus on the problem under investigation, eliminating unexpected effects from structure complexity. Experimental data taken from Esaki [3.19] was used for the data fitting.

Figure 3.5 is taken from [3.19]. The current from the reverse biased Ge p-n junction is the reverse tunneling current. This *I-V* curve is used to fit the Ge tunneling mass with the non-local band-to-band tunneling algorithm.

A 1D Ge diode structure was constructed for the tunneling mass fitting. Doping concentration was set to be the same as that reported in Esaki's paper [3.19]. The mesh and doping information is as shown in Figure 3.6. The mesh is intentionally densified near the junction position, in order to capture the band-to-band tunneling effectively.



Figure 3.5. Semi-log plots of current-voltage characteristics at 200, 300, and 350 K for a Ge p-n diode, where  $N_A \sim 2.4 \times 10^{18} \text{ cm}^{-3}$  and  $N_D \sim 10^{19} \text{ cm}^{-3}$  (taken from Fig. 1 of Ref. [3.19]).



Figure 3.6. The 1D diode structure used for Ge tunneling mass fitting.

The doping profile is assumed to be abrupt. The diode area is set to be  $0.1\mu$ m×100 $\mu$ m. The following equation is used for the current fitting,

$$I_{\exp} = D \times I_{simulation}$$
, Equation 3.7

Since the dimension of the experimental diode is not given in the experimental results, an area ratio parameter D is used to fit the current level, where D can be calculated as follows,

$$D = A_{\text{experimental}} / A_{\text{Simulation}},$$
 Equation 3.8

The value of the tunneling mass directly affects the shape of the *I-V* curve. Different values of effective tunneling mass  $m_r$  were set in the simulator. Then the simulator was re-configured and compiled. The 1D diode structure was simulated with different tunneling mass values. *I-V* characteristics was captured and compared with the experimental value, in order to find a reasonable fit.

It can be seen that reasonably good fitting in both linear and log scale can be achieved with  $m_r$  from  $0.01m_0$  to  $0.03m_0$ , up to more than one order of magnitude of current level, as seen in log-linear scale of the *I-V* curves (Fig. 3.7). The discrepancy might be due to the fact that the experimental doping data was not accurately determined, or the doping concentration and doping profile used in the simulation do not match the experimental conditions.



Figure 3.7. The fitting of simulated diode *I-V* curves (colored) with the experimental data (plotted as black line) in log-linear (top) and linear-linear (bottom) scale. Simulation data with different tunneling masses is shown on the same figure.

Besides fitting of simulation results to experimental data, a literature search on analytical modeling of band-to-band tunneling in Ge was done and compared with the results of fitting, in order to verify that the fit is reasonably accurate and the tunneling mass can be used in subsequent device simulations.

Some theoretical calculations were done by Butcher *et al.* [3.20]. By taking  $E_g$  = 0.80 eV, for electrons tunnel into the (000) minimum, the following formula can be used to calculate the tunneling mass.

$$\frac{1}{m_r} = \frac{1}{m_{lh}} + \frac{1}{m_c},$$
 Equation 3.9

where  $m_{lh} = 0.043 \ m_0$  is the light-hole mass and  $m_c = 0.036 \ m_0$  is the (000) conduction band mass.  $m_0$  is the free electron mass. The resulting tunneling mass was calculated to be  $m_r = 0.01959 \ m_0$  for Ge.

For tunneling in the (100) direction to the (111) minima, taking  $E_g = 0.66$  eV, the tunneling mass can be calculated as

$$\frac{1}{m_r} = \frac{1}{m_{lh}} + \frac{1}{3m_L} + \frac{2}{3m_T},$$
 Equation 3.10

where  $m_L = 1.6 m_0$  and  $m_T = 0.082 m_0$  are the longitudinal and transverse masses of the (111) minima. Based on this formula, the tunneling mass was calculated to be 0.03165  $m_0$ .

Hall quoted a reduced tunneling mass of  $0.02 m_0$  for band-to-band tunneling in Ge [3.21], while Ghynoweth *et al.* suggested a reduced tunneling mass of  $0.05 m_0$  [3.22].

Based on the above mentioned published theoretical and experimental data, the simulation fitting is believed to be reasonable and a tunneling mass of  $m_r = 0.02$  $m_0$  is selected for the use in Ge tunneling FET simulation.

With a more physical and accurate simulation algorithm developed, the key parameters related to band-to-band tunneling known, variability study of a Ge tunneling FET device is studied through a simulation approach. It will be discussed in the following section.

## 3.2.2 Variability Study of a Ge Tunneling FET

Process induced parameter variability is inherent in device fabrication due to process non-uniformity, and there are a few key parameters which might affect tunneling FET operation, including EOT, tunneling junction position, doping variation, etc. EOT variation and source/drain (S/D) doping profile variation is considered important in conventional CMOS devices. Tunneling FET shares a similar MOS structure as conventional MOSFETs. Thus, variation in tunneling FET performance due to variations in device parameters would be worthy of investigation. Due to the fact that tunneling FET device operation relies on gate-controlled band-toband tunneling occurring at the source-to-channel junction, the source-to-channel junction position relative to the gate edge is considered as a key parameter as well. It directly affects the gate control on the tunneling junction.

In order to measure a tunneling FET performance, a few key performance parameters are extracted and compared. The on-state current,  $I_{on}$ , threshold voltage,  $V_{th}$ , Subthreshold Swing, S, are the important quantities directly reflecting device performance.

 $I_{on}$  is taken at the supply voltage, where the gate voltage  $V_g$  equals to the supply voltage  $V_{dd}$  in a n-type tunneling FET. The off-state leakage current is taken at  $V_g = 0$  V, with drain biased at the supply voltage.

Threshold voltage is probably one of the most important parameters for a solid state switch, as it represents the boundary between the on and off states. In conventional CMOSFET, there is a physical definition of threshold voltage. It is typically defined as the gate voltage at the onset of strong inversion. However, tunneling FET only started attracting research interests in recent years. Commonly accepted definition for tunneling FET threshold voltage is not available. Boucart *et al.* attempted to define tunneling FET threshold voltage and drain bias [3.10], [3.23]. Two threshold voltages were defined according to the transition of drain current from quasi-exponential to a linear dependence with respect to applied voltages. However, this definition is complicated and the value is not easy to obtain. Constant current method is a commonly used method in extracting threshold voltage in conventional

CMOSFETs. The threshold voltage is taken to be the gate voltage at which the drain current reaches a certain fixed level. This definition is simple and the threshold voltage is easy to extract. For simplicity, constant current definition of threshold voltage would be used for this study.

Recall that the subthreshold threshold swing (S) is defined as the gate-voltage needed to induce a drain current change of one order of magnitude, in units of mV/decade. S can be expressed using the following formula,

$$S = \frac{dV_g}{d\log I_d},$$
 Equation 3.11

which can be viewed as the reciprocal of the tangent on the log-linear  $I_d$ - $V_g$  curve.

For the definition of subthreshold swing, two different definitions had been used in tunneling FET study, namely "point *S*" and "average *S*" ( $S_{ave}$ ). Point *S* is taken at a particular gate bias, which is to say at a point on the  $I_d$ - $V_g$  curve. It is defined as the reciprocal value of the tangent line on the  $I_d$ - $V_g$  curve at a given  $V_g$ . However, it has to be noted that such definition of point *S* does not reflect the on-off transition of a device where the current level changes over a few orders of magnitude. A more physical and reasonable definition of threshold voltage is the average *S*. In such a definition, a few orders of current change from below the threshold current level was considered. The average *S* was taken over this current/voltage range. This average *S* reflects the device on/off transition more specifically. The two definitions are depicted as shown in Figure 3.8.



Figure 3.8. Graphical depiction of the two definition of sub-threshold swing, point S (calculated from blue dashed line) and average S (calculated from red solid line). It can be seen that the blue dashed line has a steeper slope, which yields a smaller S value, as compared with the average S value obtained from the red solid line.

Judging from the published tunneling FET  $I_d$ - $V_g$  curves, including both simulation and experimental data, tunneling FET point *S* value reduces with increasing gate voltage. The smallest point *S* typically appears at very low drain current. The published *S* values were typically the point *S* taken at a low gate bias. Such *S* values cannot be used to judge the device performance, as it cannot be used to calculate the relationship between on-off state current and gate voltage bias. Average *S*, denoted by  $S_{ave}$ , provides a more accurate relationship. In this study, the average *S* value will be used.

### 3.2.2.1 The base structure

The base structure used in the simulation is a single gate Ge tunneling FET. Ge substrate with a body thickness ( $t_{body}$ ) of 25 nm is chosen as the base. Gate work function is chosen to be 3.9 eV, such that a reasonable threshold voltage can be obtained. Gate length of the base structure is set to 50 nm. The gate oxide thickness is chosen to be 12 Å, according to the ITRS (International Technology Roadmap for Semiconductors) roadmap for the low standby power requirement in the near future.



Figure 3.9. The base structure used for the variability study.

Source/Drain doping is assumed to be abrupt or box-like, with a p<sup>+</sup> doping concentration of  $1 \times 10^{20}$  cm<sup>-3</sup> for the source, and n<sup>+</sup> doping concentration of  $1 \times 10^{19}$  cm<sup>-3</sup> for the drain. The substrate doping is p-type, with a doping concentration of  $1 \times 10^{16}$  cm<sup>-3</sup>. The Source/Junction channel to gate edge misalignment,  $L_{ov}$  is set to be zero for the base structure. The structure is depicted as shown in Figure 3.9. A typical simulation mesh can be found in Figure 3.1.



Figure 3.10.  $I_d$ - $V_g$  characteristics and the performance parameters extracted for the base structure.

All simulations were carried out at room temperature of 300 K, with a drain bias of 0.8 V. The reference base Ge tunneling FET device is simulated by the nonlocal band-to-band tunneling algorithm developed. Device performance parameters were extracted. The *I-V* characteristics and the performance parameters under investigation are as shown in Figure 3.10. Constant current method was used to extract the threshold voltage. A current level of  $1 \times 10^{-4}$  mA/µm was chosen to be the threshold current level, which is around 6 orders of magnitude higher than that of the off-state current level. The off-state current is taken at  $V_d = 0.8$  V and  $V_g = 0$  V. A threshold voltage of about 0.2 V was extracted using this method. The on-state current is extracted at the bias of  $V_g = V_d = 0.8$  V, an on-state current of 0.223 mA/µm was obtained. From the *I-V* characteristics, we observe that the *S* value decreases with increasing gate bias. The average *S* value was extracted by averaging over 3 orders of drain current change, between a current level range of  $1 \times 10^{-7}$  mA/µm and  $1 \times 10^{-4}$  mA/µm, taking note of the fact that the threshold voltage was extracted at drain

category		V <sub>th</sub> (V)	I <sub>on</sub> (mA/μm)	$I_{off}(pA/\mu m)$	<i>S<sub>ave</sub></i> (mV/decade)
Base structure		0.157	0.045	57.00	49.8
Source Doping	10%	0.156	0.044	84.43	51.3
	-10%	0.157	0.044	57.04	49.8
Drain Doping	10%	0.156	0.044	76.57	50.9
	-10%	0.156	0.044	70.00	50.6
Channel Doping	10%	0.156	0.044	72.78	50.8
	-10%	0.156	0.044	73.95	50.8

Table 3.1. The effect of varying doping concentration in the source, drain, and channel region of a tunneling FET on its performance.

current of  $1 \times 10^{-4}$  mA/µm. Although the average *S* value extracted from this device is in sub-60 mV/decade range, the on-state current of this device is still much lower than that of the ITRS roadmap requirement, which might be a big challenge in the application of tunneling FET devices.

Some simulations regarding the doping concentration variation were done in a preliminary study. The doping concentration can be controlled with a negligible degree of variation, by adopting advanced implantation or epitaxy techniques. The effect of up to 10% variation in the source and drain doping concentration were considered. Doping concentration induced device performance variation is negligible. The results are as shown in Table 3.1.

It is observed that very small device performance variation was caused by the doping concentration variation. This is due to the fact that band-to-band tunneling only occurs at significantly high electric field. The built-in source-to-channel and drain-to-channel electric field caused by doping is orders of magnitude lower than when band-to-band tunneling occurs. Band-to-band tunneling current is mainly

limited by the tunneling barrier width, for a given material system. The tunneling barrier width is related to the electric field at the junction. For a  $p^+p$  junction with  $p^+$  doping of  $1 \times 10^{20}$  cm<sup>-3</sup> and p doping of  $1 \times 10^{16}$  cm<sup>-3</sup>, varying the  $p^+$  doping concentration by  $\pm 10\%$  only causes less than 1% of electric field variation (from  $3.53 \times 10^6$  V/m to  $3.52 \times 10^6$  V/m and  $3.54 \times 10^6$  V/m ). Thus, the effect of doping concentration variation is negligible for a tunneling FET.

# 3.2.2.2 EOT variation

Gate dielectric thickness is one of the most important parameters for MOS structures, as gate dielectric directly affects the potential coupling between the gate and the channel. EOT of the base structure is set to be 12Å, a  $\pm$ 10% variation was assumed in the study. Devices with equivalent oxide thickness of 90%, 95% and 105%, 110% of the initial value were simulated. Device performance parameters are extracted and compared with the initial reference value, focusing on the change of threshold voltage and on-state current, as well as the *S* value.

The main conduction mechanism of the tunneling FET device is band-toband tunneling. Current conduction starts when the conduction band of the channel is pulled below the valence band of the source side. It is directly controlled by the channel potential, with the source tied to ground potential. As in all MOS structures, dielectric thickness of the device determines the amount of potential drop across the dielectric layer, which in turn determines the potential of the channel given a particular gate bias. The potential drop on the dielectric layer can be derived by Gauss' law:

$$E_{channel} = -\frac{Q_{channel}}{\varepsilon_{channel}}$$
 Equation 3.12

$$\varepsilon_{ox}E_{ox} = \varepsilon_{channel}E_{channel}$$
, Equation 3.13

We can obtain the potential drop across the dielectric,

$$V_{ox} = t_{ox} E_{ox} = \frac{t_{ox}}{\varepsilon_{ox}} \bullet (-Q_{channel}),$$
 Equation 3.14

where *E* is the electric field,  $\varepsilon$  is the dielectric constant,  $Q_{channel}$  is the total channel charge,  $t_{ox}$  is the dielectric thickness, and  $V_{ox}$  is the potential drop across the gate dielectric. For a MOS device with thicker gate dielectric, in order to achieve same amount of potential change on the substrate, a larger gate bias is required, as the potential drop across the gate dielectric is larger comparing to a device with thinner gate dielectric.

This is also applicable on the threshold voltage derivation of a tunneling FET. A higher threshold voltage is required for a device with larger dielectric thickness.

As it can be seen from Figure 3.11 that the threshold voltage is directly affected by the EOT variation, around  $\pm 8\%$  of threshold voltage variation can be induced by 10% of EOT change, and up to around  $\pm 18$  mV of threshold voltage variation is observed.



Figure 3.11. The effect of a change in EOT on the (top) the *I-V* characteristics, (middle) threshold voltage, and (bottom) percentage of threshold voltage change in a Ge tunneling FET device.

One of the most obvious effects of the threshold voltage variation can be seen from the on-state current variation. Significant on-state current variation is caused by the EOT variation. Due to the change of threshold voltage caused by the EOT variation, the effective on-state gate overdrive is also changed. More than 20% of the on-state current increase was observed if the EOT of the device is reduced by 10%. Considering the fact that EOT of the base Ge tunneling FET structure is only 12 Å, a 1.2 Å EOT variation can cause more than 20% of drive current change. Such amount of drive current variation is not desirable, and a very tight process control is required to limit the EOT thickness variation. Sub-1 Å EOT variation is needed to lower the drive current variation for this base device structure, in order to keep the drive current variation below 10%. This could impose a serious challenge in the process control of tunneling FETs.

From Figure 3.12, we can see that the reduction of EOT thickness causes a larger change in current as compared with the increase of EOT thickness. Reduction of 10% of EOT thickness caused more than 25% of drive current increase while less than 20% of drive current reduction was caused by 10% increase of the EOT thickness. This means that effect of EOT variation could be even more severe for devices with thinner EOT thickness.



Figure 3.12. Effect of EOT variation on the on-state current of the Ge tunneling FET in absolute value (top) and relative value (bottom).

Average *S* values were also extracted and plotted, as shown in Fig. 3.13. As mentioned in previous sections, the average *S* of the Ge tunneling FET was extracted by averaging the drain current over three orders of magnitude below the threshold voltage. Sub-60 mV/decade of average *S* value was observed for the base structure as well as the device with  $\pm 10\%$  of EOT variation. The observed average *S* values were in the regime of 50 ~ 60 mV/decade. It is interesting to note that the average *S* values



Figure 3.13. Effect of EOT variation on the average subthreshold of a Ge tunneling FET.

tend to saturate in negative EOT variation. Recall that band-to-band tunneling generation rate is modeled as follows,

$$dG_{BTBT} = \frac{qm_{dos}^*kT}{2\hbar^3} \frac{n_c p_v - n_i^2}{\left(n_c + n_i\right)\left(p_v + p_i\right)} W \bullet T_{tun} \bullet dE, \qquad \text{Equation 3.15}$$

The band-to-band tunneling current is solely determined by the tunneling rate with a given device. All other parameters are fixed once the device material system is fixed. Band-to-band tunneling rate is determined by the change of tunneling barrier thickness, which is in turn controlled by gate voltage. Gate control of the tunneling barrier thickness is obtained through the potential coupling across the gate dielectric. By decreasing the EOT of the device, gate voltage is more directly coupled with the channel potential. When EOT of a tunneling FET device decreases, band-to-band tunneling rate approaches a fixed value, which is solely determined by the Fermi distribution over both sides of the tunneling barrier. Drive current at each bias would also approach a particular value if there is no voltage drop across the dielectric layer. In the extreme case where EOT is zero, a fixed *I-V* characteristic should be obtained, thus the average *S* would be a fixed value.

This indicates that for a given tunneling FET device, the *S* value would also be limited by the material system and operation conditions. Although tunneling FET has the potential of achieving sub-60 mV/decade *S* values, the subthreshold swing cannot decrease to below a fixed value either.

#### 3.2.2.3 Junction position variation

In tunneling FET devices, band-to-band tunneling occurs at the source to channel junction. Control of the tunneling current is obtained by gate voltage modulation. The position of the tunneling junction relative to the gate edge is an important parameter for tunneling FETs. The impact of the variation of the junction position was studied and documented in this section. Abrupt source doping profile is assumed, such that a clear source-to-channel junction position could be identified. The junction to gate edge relative position is denoted by  $L_{ov}$ , as indicated in Figure 3.9. Moving the junction position further away from the gate edge is denoted by a negative  $L_{ov}$ , while moving the junction position towards under the gate is denoted by a positive  $L_{ov}$ .



Figure 3.14. The effect of junction position variation on the threshold voltage variation in absolute value (top) and relative percentage (bottom).

Junction position variation up to  $\pm 10\%$  of the gate length was considered. By setting the junction position from 5 nm away from the gate edge to 5 nm under the gate, device *I-V* characteristics were simulated, and performance parameters were extracted and compared.

As one of the most important parameters for a solid state switch, threshold voltage of devices with different  $L_{ov}$  was extracted and plotted in Figure 3.14. From Figure 3.14, it can be seen that the threshold voltage of the device increased significantly if the junction position was moved away from the gate edge. More than 40 mV of threshold voltage increase was observed if the junction position was moved 5 nm away from the gate edge. This amount of threshold voltage change can be attributed to the decreased electric field at the tunneling junction. By moving the junction position further away from the gate edge, effectively an intrinsic region is inserted between the junction and the gate edge, and a potential drop on this intrinsic region is decreased. Larger gate bias is required in order to build up same amount of electric field at the tunneling junction, allowing same amount of tunneling current. This proposed model is shown in the Figure 3.15 (top).

In order to justify this model, electric field distribution along the gate direction was extracted for two devices, at the same bias voltage near the threshold voltage. The electric field of the base structure as well as the device with  $L_{ov} = -5$  nm was extracted and plotted for comparison, as shown in Figure 3.15. For both devices, the gate edge is located at 0.07 µm. It can be observed that the peak electric field appears near the gate edge for both devices. However, the peak electric field of the



Figure 3.15. The device model in explaining the effect of moving the junction away from the gate edge (top) and the electric field projection on the gate direction for two devices with  $L_{ov} = 0$  (blue) and  $L_{ov} = -5$  nm (red) at a gate bias of 0.2 V (bottom).

base structure (blue) is significantly larger than when the source junction is moved away from the gate edge (red). Peak electric field of  $1.0216 \times 10^7$  V/m is observed for the base structure, while the peak electric field is found to be only  $6.3818 \times 10^6$  V/m for the device with  $L_{ov}$ = -5 nm. A shoulder electric peak can also be observed for the device with  $L_{ov}$ = -5 nm, located between x = 0.065 µm and the gate edge (0.07 µm). This location corresponds to the intrinsic region between the source/channel junction to the gate edge. Electric field with an average value around  $1 \times 10^6$  V/m was observed in this region. Horizontal electric field is partially contributed by the Drain to Source bias, while the vertical component is mainly contributed by the gate. This shoulder electric field caused a small amount of potential drop, which in turn reduced the effective potential drop on the tunneling junction.

The extracted electric field distribution fully supported the model formulated. The intrinsic region inserted between the gate edge and source/channel junction reduced the peak electric field and thus reduced the tunneling current, causing an increase of the threshold voltage.

While the source/channel junction is shifted towards under the gate edge, a smaller threshold voltage is observed. This is due to the increase of the horizontal electric field along the channel induced by the reduction of the effective physical gate length. Here, the effective physical gate length is defined as the distance between source/channel junction and drain/channel junction. This threshold voltage variation is expected to be more severe for devices with even shorter gate length.

Besides the effect on threshold voltage, junction position variation affects the on-state current significantly as well.

The on-state current is as shown in Figure 3.16 for devices with different source/channel to gate misalignment. Significant on-state current drop up to more than 50% is observed for  $L_{ov} = -5$  nm, a current drop of around 15% is observed for a 2 nm misalignment. Effects of moving the junction position towards under the gate is relatively less obvious, a maximum drive current increase of around 10% is observed for  $L_{ov} = 5$  nm.

Stringent implantation and annealing condition control is required to control the junction position, since junction position misalignment in nm-scale would induce a few tens percent of current variation.



Figure 3.16. The extracted on-state current for devices with different source/channel junction to gate edge misalignment (upper) and the relative percentage variation with respect to the base structure whose  $L_{ov}$  is zero(lower).



Figure 3.17. Effect of junction misalignment on the average S value.

Average S values were extracted and plotted in Figure 3.17. The S value variation further emphasized the importance of controlling the junction position variation. By moving the junction away from the gate edge, a portion of the gate voltage was dropped across the intrinsic region. The gate control over the tunneling junction became weaker, thus the average S value was reduced. When the junction was moved towards under the gate, the lateral electric field caused by source-drain bias became more prominent, and effectively the gate control was weakened, causing the S value to increase.

The junction position variation simulation indicated that in order to fabricate tunneling FETs with consistent device performance, stringent process condition control and optimization is required.

# 3.3 Summary

In this chapter, a non-local band-to-band tunneling simulation algorithm was developed. Its application in tunneling FET simulation was discussed. Simulation approach was taken to study the effect of process related variations on the device performance of a single gate Ge tunneling FET. Effects of key process variation on threshold voltage, on-state current as well as the average subthreshold swing are studied. The simulation results indicate that tunneling FET device performance can be greatly affected by process variations. Tight process control is therefore required in order to fabricate tunneling FET device with consistent and tightly distributed performance parameters. Sub-1 Å EOT variation is required to limit the variation of on-state current within 10%. Sub-nm scale junction to gate edge alignment control is expected. These requirements might be potential drawbacks for the widespread adoption of tunneling FET devices.

## **References:**

- [3.1] E. O. Kane, "Zener Tunneling In Semiconductors," J. Phy. Chem. Solids, vol. 12, July 1959 1959.
- [3.2] E. O. Kane, "Theory of Tunneling," *Journal of Applied Physics*, vol. 32, p. 83, 1961.
- [3.3] P.-F. Wang, T. Nirschl, D. Schmitt-Landsiedel and W. Hansch., "Simulation of the Esaki-tunneling FET," *Solid-State Electronics*, vol. 47, pp. 1187-1192, 2003.
- [3.4] E.-H. Toh, G-H Wang, G. Samudra and Y.-C. Yeo, "Device physics and design of germanium tunneling field-effect transistor with source and drain engineering for low power and high performance applications," *Journal of Applied Physics*, vol. 103, 2008.
- [3.5] E.-H. Toh, G-H Wang, G. Samudra and Y.-C. Yeo, "Device physics and design of double-gate tunneling field-effect transistor by silicon film thickness optimization," *Applied Physics Letters*, vol. 90, pp. 263507-1, 2007.
- [3.6] E.-H. Toh, G-H Wang, G. Samudra and Y.-C. Yeo, "Device physics and guiding principles for the design of double-gate tunneling field effect transistor with silicon-germanium source heterojunction," *Applied Physics Letters*, vol. 91, 2007.
- [3.7] K. Boucart and A. M. Ionescu, "A new definition of threshold voltage in tunnel FETs," *Solid-State Electronics*, vol. 52, pp. 1318-23, 2008.
- [3.8] K. Boucart and A. M. Ionescu, "Double-gate tunnel FET with high-k gate dielectric," *IEEE Transactions on Electron Devices*, vol. 54, pp. 1725-33, 2007.
- [3.9] K. Boucart and A. M. Ionescu, "Double gate tunnel FET with ultrathin silicon body and high-k gate dielectric," Montreux, Switzerland, 2007, pp. 383-386.
- [3.10] K. Boucart and A. M. Ionescu, "Threshold voltage in tunnel FETs: physical definition, extraction, scaling and impact on IC design," Piscataway, NJ, USA, 2007, pp. 299-302.
- [3.11] Medici, "User Guide 2007\_12 version."
- [3.12] G. A. M. Hurkx, D.B.M. Klaassen, M. P. G.Knuvers, "A novel compact model description of reverse-biased diode characteristics including tunnelling," Bristol, UK, 1990, pp. 49-52.
- [3.13] C. Shen, L. Yang E.-H. Toh, C.-H. Heng, G. Samudra and Y.-C. Yeo, "A new robust non-local algorithm for band-to-band tunneling simulation and its application to tunnel-FET," presented at the International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA), Hsinchu, Taiwan, 2009.
- [3.14] GSS, General-purpose Semiconductor Semulator, an open source TCAD software, it enables two-dimensional numerical simulation of semiconductor device with well-known drift-diffusion and hydrodynamic method. More can be found on "http://gss-tcad.sourceforge.net/".
- [3.15] Congenda, a company specialized in developing TCAD simulators and related technical services, more can be found on "<u>http://www.cogenda.com/.</u>"
- [3.16] Z. H. Huang, T. E. Feuchtwang, P. H. Cutler and E. Kazes, "Wentzel-Kramers-Brillouin method in multidimensional tunneling," *Physical Review A* (*Atomic, Molecular, and Optical Physics*), vol. 41, pp. 32-41, 1990.

- [3.17] P. J. Price and J. M. Radcliffe, "Esaki Tunneling," *IBM Journal of Research and Development*, vol. 3, 1959.
- [3.18] B. C. Paul, *et al.*, "Impact of a process variation on nanowire and nanotube device performance," *IEEE Transactions on Electron Devices*, vol. 54, pp. 2369-76, 2007.
- [3.19] L. Esaki, "Discovery of the tunnel diode," *IEEE Transactions on Electron Devices*, vol. ED-23, pp. 644-7, 1976.
- [3.20] P. N. Butcher, K. F. Hulme and J. R. Morgan, "Dependence of peak current density on acceptor concentration in germanium tunnel diodes," *Solid-State Electronics*, vol. 5, pp. 358, in27-in28, 359-360.
- [3.21] R. N. Hall, "Excess and Hump Current in Esaki Diodes", Bulletin of the American Physical Society, Vol 5. pp38, 1960.
- [3.22] A. G. Chynoweth, W. L. Feldmann and R.A. Logan, "Excess Tunnel Current in Silicon Esaki Junctions," *Physical Review*, vol. 121, p. 684, 1961.
- [3.23] K. Boucart and A. M. Ionescu, "A new definition of threshold voltage in Tunnel FETs," *Solid-State Electronics*, vol. 52, pp. 1318-1323, 2008.
## Chapter 4

# Conclusion and Future Trends of Tunneling FET Research

MOSFET is approaching its fundamental limits. It is believed that new device concepts/structures would emerge and find their application in post-CMOS era. Tunneling Field Effect Transistor (tunneling FET, or TFET) has been under active research in recent years. As one of the most promising device concepts for the post-CMOS era, it would attract more research interest in the coming years.

W Y Choi *et al.* claimed that they had successfully demonstrated an nchannel tunneling FET with 52 mV/decade subthreshold swing in 2007 [4.1]. This is probably the first published experimental results for silicon TFETs with sub 60 mV/decade swing. Conventional SiO<sub>2</sub> dielectric with Si substrate was used. It is believed that by adopting advanced fabrication technique and new material, device performance could be further enhanced. In experimental efforts of this study, experimental challenges in tunneling FET fabrication are first discussed. In order to fabricate tunneling FET device with good device performance, integration of advanced fabrication technology is required. Experimental efforts in fabrication of tunneling FET device were carried out. Application of high- $\kappa$  gate dielectric enables low gate leakage; HfO<sub>2</sub> grown through Atomic Layer Deposition (ALD) was used in the fabrication process. Advanced doping technique is capable of forming abrupt junction profile and enhance band-to-band tunneling rate; the technique of dopant segregation was studied and applied in the fabrication process, it is supposed to be useful in forming abrupt doping profile. New materials with low band gap could enhance band to band tunneling; low band gap material Ge/SiGe was applied in the study. Integration of those techniques would help in realization of working tunneling FET devices. Working tunneling FET devices were fabricated; device characteristics were measured and discussed.

From simulation approach, comprehensive study has been carried out by Toh *et al.*[4.2]-[4.4], Patel *et al.*[4.5] and more [4.6]-[4.7] in the past few years. Some device design rules and characteristics were discussed. However, all the simulation studies were carried out with commercially available device simulators, whose accuracy in band-to-band tunneling simulation is questionable. The formulation of the band-to-band tunneling algorithm is not physical in most of the simulators used. In this work, a novel non-local band-to-band tunneling algorithm is developed and tested, based on a 2D open source general purpose semiconductor simulator. This algorithm is found to be physical and robust in band-to-band tunneling simulation.

Further on, simulation study of tunneling FET was carried out using the nonlocal band to band tunneling algorithm developed. Process induced parameter variation was studied. It is found that the doping concentration variation would only induce negligible performance variation in tunneling FETs. Equivalent oxide thickness (EOT) variation was found to affect tunneling FET performance significantly. Comparing with the base structure, whose EOT is 12 Å, a 10% EOT variation could induce up to more than 20% variation of the on-state current. This result indicated that tight EOT control is required. It is also found that if the tunneling junction is mis-aligned away from the gate; significant on-state current reduction is expected. Stringent doping profile control is required in the fabrication process.

Slightly after this study, some more experimental results were published. F. Mayer *et al.* discussed the impact of SOI, Si<sub>1-x</sub>Ge<sub>x</sub>OI and GeOI substrates on the tunneling FET performance, and claimed a 42 mV/decade p-channel SOI tunneling FET [4.9]. T. Krishnamohan *et al.* reported double-gated strained Ge heterostructure tunneling FET with record high drive current and sub 60 mV/decade subthreshold slope [4.10]. These experimental results indicated the potential of tunneling FET devices in future applications. However, it is also true that the reported sub 60 mV/decade subthreshold swing is usually taken from a point of the  $I_d$ - $V_{gs}$  curve, which means these values did not reflect the true device performance throughout the on-off transition. What is more important is that the on-state current of the reported devices is usually very low, typically in sub  $\mu$ A/ $\mu$ m regime. The record high on-state current reported by T. Krishnamohan *et al.* was taken at 3 V to 4 V supply/gate voltage, which is not a true reflection of the device performance, as the supply voltage is much higher than the reasonable supply voltage for modern integrated circuits.

The reported experimental results indicated that further process development and optimization is required in realizing tunneling FET devices with good performance characteristics. Fabrication process of lateral tunneling FET is not selfaligned, separate source/drain masking is required in the implantation process. It is expected that there would be some difficulties in lithography process for devices with smaller gate length. There would be a need for the development of new tunneling FET structures which provide better process control and easier integration with conventional CMOSFET devices.

In conclusion, tunneling FET is a potential device concept for post-CMOS era, due to its low leakage current and small subthreshold swing. However, very stringent process requirement and variation control are required in order to fabricate such devices successfully. Experimental exploration is still very challenging, as advanced process technology and very tight process control are required. Structure innovation, advanced material study could help in realizing tunneling FET devices. Material systems like Ge/Si hetero junction or group III-V compound materials might find their application in tunneling FET fabrication.

#### Reference

- [4.1] W. Y. Choi, *et al.*, "Tunneling field-effect transistors (TFETs) with subthreshold swing (SS) less than 60 mV/dec," *IEEE Electron Device Letters*, vol. 28, pp. 743-5, 2007.
- [4.2] E.-H. Toh, G. H. Wang G. Samudra and Y-C. Yeo, "Device physics and guiding principles for the design of double-gate tunneling field effect transistor with silicon-germanium source heterojunction," *Applied Physics Letters*, vol. 91, pp. 243505-1, 2007.
- [4.3] E.-H. Toh, G. H. Wang G. Samudra and Y-C. Yeo, "Device design and scalability of a double-gate tunneling field-effect transistor with silicon germanium source," *Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers*, vol. 47, pp. 2593-2597, 2008.
- [4.4] E.-H. Toh, G. H. Wang G. Samudra and Y-C. Yeo, "Device physics and design of germanium tunneling field-effect transistor with source and drain engineering for low power and high performance applications," *Journal of Applied Physics*, vol. 103, 2008.
- [4.5] N. B. Patel, A. Ramesha and S. Mahapatra, "Performance enhancement of the tunnel field effect transistor using a SiGe source," Mumbai, India, 2007, pp. 111-114.
- [4.6] K. K. Bhuwalka, et al., "Vertical tunnel field-effect transistor with bandgap modulation and workfunction engineering," Leuven, Belgium, 2004, pp. 241-244.
- [4.7] K. Boucart and A. M. Ionescu, "Double-gate tunnel FET with high-k gate dielectric," *IEEE Transactions on Electron Devices*, vol. 54, pp. 1725-33, 2007.
- [4.8] A. S. Verhulst, W. G. Vaenbegh, K. Maex, and G. Groeseneken, "Boosting the on-current of a n -channel nanowire tunnel field-effect transistor by source material optimization," *Journal of Applied Physics*, vol. 104, 2008.
- [4.9] F. Mayer, et al., "Impact of SOI, Si1-xGexOI and GeOI substrates on CMOS compatible tunnel FET performance," in 2008 IEEE International Electron Devices Meeting, IEDM 2008, December 15, 2008 - December 17, 2008, San Francisco, CA, United states, 2008.
- [4.10] T. Krishnamohan, D. Kim, S. Raghunathan and K. Saraswat, "Double-gate strained-Ge heterostructure tunneling FET (TFET) With record high drive currents and 60 mV/dec subthreshold slope," Piscataway, NJ, USA, 2008, p. 3 pp.

# Appendix A

### List of publication

- R. T.-P. Lee, L. Yang, K.-W. Ang, T.-Y. Liow, K.-M. Tan, A. S.-W. Wong, G. S. Samudra, D. Z. Chi, and Y.-C. Yeo, "Material and electrical characterization of nickel silicide-carbon as contact metal to silicon-carbon source and drain stressors," *Materials Research Society Spring 2007 Meeting*, San Francisco, CA, Apr. 9-13, 2007.
- R. T. P. Lee, L.-T. Yang, T.-Y. Liow, K.-M. Tan, A. E.-J. Lim, K.-W. Ang, D. M. Y. Lai, K. M. Hoe, G.-Q. Lo, G. S. Samudra, D. Z. Chi, and Y.-C. Yeo, "Nickel-silicide:carbon contact technology for n-channel MOSFETs with silicon-carbon source/drain," *IEEE Electron Device Letters*, vol. 29, no. 1, pp. 89-92, Jan. 2008.
- C. Shen, L. T. Yang, E.-H. Toh, C.-H. Heng, G. S. Samudra, and Y.-C. Yeo, "A new robust non-local algorithm for band-to-band tunneling simulation and its application to tunnel-FET," *International Symposium on VLSI Technology, Systems and Applications (VLSI-TSA)*, Hsinchu, Taiwan, Apr. 27-29, 2009, pp. 113-114.
- C. Shen, L. T. Yang, G. S. Samudra, and Y.-C. Yeo, "A new robust non-local algorithm for band-to-band tunneling simulation and its application to tunnel-FET," *Solid-State Electronics*, vol. 57, pp. 23-30.
- P.-F. Guo, L.-T. Yang, Y. Yang, L. Fan, G.-Q. Han, G. Samudra, and Y.-C. Yeo, "Tunneling field effect transistor: Effect of strain and temperature on tunneling current," *IEEE Electron Device Letters*, vol. 30, 2009.
- L. Fan, L.-T. Yang, Y. Yang, P.-F. Guo, G. Samudra, and Y.-C. Yeo, "A non-local algorithm for simulation of band-to-band tunneling in a heterostructure tunnel field-effect transistor (TFET)," *Extended Abstracts of the 2009 International Conference on Solid State Devices and Materials*, Sendai, Japan, Oct. 7-9, 2009.
- Y. Yang, X. Tong, L. Yang, P. Guo, L. Fan, G. S. Samudra, and Y.-C. Yeo, "Capacitances in tunneling field-effect transistors," *Extended Abstracts of the 2009 International Conference on Solid State Devices and Materials*, Sendai, Japan, Oct. 7-9, 2009.