12-12-2017

Optimizing Fabrication and Modeling of Quantum Dot Superlattice for FETs and Nonvolatile Memories

Pial Mirdha
University of Connecticut - Storrs, pialmirdha@gmail.com

Follow this and additional works at: https://opencommons.uconn.edu/dissertations

Recommended Citation
https://opencommons.uconn.edu/dissertations/1686
Quantum Dot Superlattice (QDSL) are novel structures which can be applied to transistors and memory devices to produce unique current voltage characteristics.

QDSL are made of Silicon and Germanium with an inner intrinsic layer surrounded by their respective oxides and in the single digit nanometer range. When used in transistors they have shown to induce 3 to 4 states for Multi-Valued Logic (MVL). When applied to memory they have been demonstrated to retain 2 bits of charge which instantly double the memory density.

For commercial application they must produce consistent and repeatable current voltage characteristics, the current QDSL structures consist of only two layers of quantum dots which is not a robust design.

This thesis demonstrates the utility of using QDSL by designing MVL circuit which consume less power while still producing higher computational speed when compared to conventional cmos based circuits.

Additionally, for reproducibility and stability of current voltage characteristics, a novel 4 layer of both single and mixed quantum dots are demonstrated. The stacking of QDSL of more than 2 layers allows greater charge storage which can add lead to more distinct MVL and memory states. This QDSL structure is verified using AFM. Also demonstrated is the capability to assemble only one layer of QDSL.

Finally a physics and surface potential based numerical model is developed which incorporates the QDSL structures charge storage. This is can be used to model transistors and memory for circuit application or for individual device physics analysis for optimization. The QDSL charge storage in modeled in a computationally less intensive way when compared to their derivation from quantum mechanics.
Optimizing Fabrication and Modeling of Quantum Dot Superlattice for FETs and Nonvolatile Memories

Pial Mirdha

B.S., University of Connecticut, 2006
M.S., University of Connecticut, 2017

A Dissertation
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy
at the
University of Connecticut

2017
APPROVAL PAGE

Doctor of Philosophy Dissertation

Optimizing Fabrication and Modeling of Quantum Dot Superlattice for FET and Nonvolatile Memory
Presented By

Pial Mirdha B.S.

Major Advisor___________________________________________
Dr. Faquir Jain

Associate Advisor_______________________________________
Dr. John Ayers

Associate Advisor_______________________________________
Dr. Lei Wang

Associate Advisor_______________________________________
Dr. Evan Heller

Associate Advisor_______________________________________
Dr. John Chandy

University of Connecticut 2017
Dedication

To my mother, Anwari ”Daisy” Nayeem, without whom this would not be possible.
Acknowledgements

I would like to thank Dr. Faquir Jain, my thesis advisor, with whom I have worked these past six years and whose guidance has been instrumental to my success.

I would also like to thank my advisory committee Dr. John Ayers, Dr. Lei Wang, Dr. John Chandy, and Dr. Evan Heller for all their help.

Additionally I would like to thank Dr. Pik-Yiu Chan, Dr. Murali Lingalugari, Dr. Evan Heller, Jon Ficket, Dr. Jun Kondo, Dr. Ernesto Suarez and Barath Parthasarathy. I have had the chance to work closely with these individuals closely and their help in running experiments, fabricating devices, developing code, advice and in many other things, have helped me in innumerable ways. I cannot thank them enough.

I would also like to thank the extended lab family of present and past researcher in Dr. Jain lab who have made themselves available for consultation and dedicated their time to help my research.

I would like to thank University of Connecticut ECE department, all the faculty, the staff, and all other graduate students with whom I have had the pleasure of working.
Contents

Approval ii

Dedication iii

Acknowledgements iv

1 Transistor Scaling Challenge 1

1.1 Scaling at an End? 1

2 SWS-FET Based Multi-Valued Logic Circuits 7

2.1 Spatially Wave-function Switched FET 7

2.1.1 SWS-FET Device Characteristics 8

2.1.2 BSIM Based SWS-FET model 10

2.2 Multiplexer and Di-multiplexer 12

2.2.1 SWS-FET Based Multiplexer and Di-multiplexer Design 13

2.2.2 Design and Results for 2:1 Multiplexer and Di-multiplexer 14

2.3 Discussion 19

3 Multi-Layered Quantum Dot Superlattice for Optimized Performance of Transistors and Memories 20

3.1 Quantum Dot Superlattice Novel Semiconductor Material 20

3.1.1 QDSL Applied to Transistors 21
3.1.2 QDSL Applied to Flash Memory ........................................ 24
3.2 QDSL Charge Storage Capability ........................................ 27
3.3 Quantum Dot Supper-Lattice Characteristics .......................... 31
  3.3.1 QDSL Fabrication ..................................................... 39
  3.3.2 QDSL Four Layer Fabrication ....................................... 42
3.4 QDSL Optimization ........................................................ 43
3.5 Discussion ..................................................................... 48

4 Surface Potential Model of QDG-FET 50
  4.1 Surface Potential Based Quantum Dot Model ......................... 50
  4.2 New Formalism for Surface Potential Model Incorporating QDSL . 59
    4.2.1 Modified Surface Potential Results for MVL Devices ......... 61
    4.2.2 Modified Surface Potential Results for QDGNVM Devices .... 63
  4.3 Discussion ..................................................................... 65

5 Conclusion 66

6 References 69
List of Figures

1.1 Increasing use of periodic elements for fabrication of transistors and circuits (1) 3

1.2 Evolution of Transistor from a) planer to 3D channel and Gate known as b) FinFET. (2) 4

2.1 Quantum Well Based SWS-FET with two Drains for MVL logic applications. 8

2.2 Quantum Simulation of SWS-FET shows electron wave function confined to each of the channels. 9

2.3 Quantum Simulation of SWS-FET shows electron wave function confined to each of the channels. 10

2.4 Two BSIM transistors are used to mimic twin Drain SWS-FET behavior using variable threshold voltage for each transistor and timing 11

2.5 a) Shows the 2:1 truth table of MUX operation while b) show CMOS based logic circuit (3; 4) 13

2.6 4:1 to conventional CMOS MUX is shown with roughly 38 transistors(4) 13

2.7 a) 2:1 Mux with an integrated CMOS and SWS-FET design and b) one with only an SWS-FET design. This SWS-FET only design can further reduce number of transistors and reduce manufacturing complexity from having to integrate CMOS and III-IV materials. 14

2.8 2:1 simulation results for both CMOS and SWS-FET integrated circuit and SWS-FET only. The output Z correspond correctly to the input S. 16
2.9 2:1 Demux simulation results for both circuits shown in Fig.2.7 with the inputs
and outputs reversed .................................................. 16

2.10 4:1 Mux and DeMux circuits is designed with a) using a combination of CMOS
and a four channel SWS-FET and b) an SWS-FET based circuit only. ........ 17

2.11 4:1 Mux simulation for both combination of CMOS and SWS-FET and SWS-
FET design only with output shown in Z. .................................... 18

2.12 4:1 DeMux simulation results for both circuits shown in Fig.2.10 with the
inputs being Z and outputs A, B, C, D ........................................ 18

3.1 Transistors with QDSL applied show significantly different $I_D - V_g$ which is
ideal for MVL application .................................................. 22

3.2 a) Channel current tunnels into QDSL layer and leaks in the Gate as a function
of Gate Voltage while b) shows the intermediate state caused when charges
are stored in QDSL layer causing the Drain current to temporarily saturate. 23

3.3 ................................................................. 24

3.4 QDSL applied to Flash memory can optimize charge retention and show multi-
bit storage ................................................................. 25

3.5 Fabricated deice structures in a) produces the $I_D - V_g$ results shown in b).
Multi-bit storage in the multiple $V_{th}$ shift seen ............................... 26

3.6 a) The structure of MOSCAP consists of the QDSL layer with only the
cladding acting as tunnel insulator and ab$HfO_2/Al_2O_3$ stack acting as a Gate
Insulator. b) Demonstrates that control MOSCAP has virtually no charge
storage capability ......................................................... 28

3.7 Si and Ge MOSCAP show hysteresis which indicates charge storage with Si
QDSL storing significantly more, this gives designers a chance to chose QDs
used based on charge storage needs ........................................ 29

3.8 Low Frequency C-V shows charges tunneling into QDSL layer, causing capac-
itance to go down temporarily ........................................... 30
3.9 Self-Assembled QD will only assemble over p-type semiconductor . . . . . . 31
3.10 a) Si QDSL will only assemble on p-type wafers in 2 layers and b) it is desired
to get to 4 layer of QDSL assembly of only Si or Ge or a combination of the two 32
3.11 a) An AFM surface profile of a line sample with Germanium QDSL assembled,
dark band represent where QDs did not assemble as seen in (b) where an
average surface profile cut is shows, this shows Germanium QD sizes to be
4 nm. A three dimension surface profile of Silicon QDSL is seen in (c) with
some sample defects and AFM data collection artifacts but an average surface
profile cut shows in (d) that Silicon QDs to be 6 nm . . . . . . . . . . . . . 33
3.12 Bandgap profile of one layer of Silicon QDSL resembles a Quantum Well band
structure . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
3.13 Silicon QDSL miniband structure ev b) Germanium QDSL miniband struc-
ture. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 35
3.14 Density of states with first three minibands filled for a Silicon based QDSL . 36
3.15 Total charge in minibands calculated for Silicon and Germanium QDSL for
an applied voltage . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37
3.16 Photoluminescence measurement shows altered bandgap of Germanium QDSL
at 0.97 eV compared to Bulk value of 1.24 eV. . . . . . . . . . . . . . . . . 38
3.17 Sonication assisted oxidation of quantum dots forming the cladding layer. (5) 40
3.18 PH of QD solution need to be below 4.9 for uniform Self-Assembly (5). . . 41
3.19 a) Fabricated QDG-FET results have non idealized intermediate states b)
compared to idealized states which are more distinct. . . . . . . . . . . . . . 43
3.20 a) Fabricated QDG-NVM has multi-bit storage capability but states are not
distinct enought b) when compared to idealized case. . . . . . . . . . . . . . 44
3.21 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 45
3.22 QDSL made with a bottom two layers of Silicon and two layers of Germanium
on top. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 45
3.23 a) QDSL made from four layer of Germanium with b) showing 3D AFM profile and c) showing a cross-section profile of heights along the wafer. Four layers can lead to increased charge storage capability.

3.24 a) Single layer of Silicon QDSL formed on a wafer with b) showing a noisy AFM image but cross-section cut c) is taken where in a region with very little noise. The single layer of QDSL is formation is predicated on significantly lower the QD density of the solution.

3.25 Transmission measurements of poly silicon wafer with and without 2 layer of Germanium QDSL. The shift in transmission shows the effect of Germanium QDSL. Additionally previously collected Silicon absorption spectra is shown.

4.1 Surface potential guess is shown in blue and the result of iteration in red.

4.2 N channel MOSFET (NMOS) basic device layout.

4.3 NMOS showing charges.

4.4 Surface Potential with Guess vs Iterative result,(b) shows a close up where iteration converges to a solution.

4.5 Total charge in minibands calculated for Silicon and Germanium QDSL for an applied voltage.

4.6 Silicon QDSL utilized in a) MVL transistor and b) in a single bit memory.

4.7 Silicon QDG transistor showing a) simulated results with intermediates step and b) actual fabricated device results.

4.8 Silicon QDG memory showing a) simulated results with threshold voltage shift and b) actual fabricated device results.
Chapter 1

Transistor Scaling Challenge

1.1 Scaling at an End?

Integrated circuits have thus far followed a predictable doubling of transistors density every two years, called Moore’s law (6). This has produced tremendous dividend where increasing computational power has been able to be delivered leading to popularization of computers and drastically shaping every field. This miracle has been largely achieved by scaling transistors ever smaller using design rules proposed by Dennard (7). In scaling transistors the channel length is shortened, the factor by which the channel is scaled then dictates the dimensions of tunnel insulator oxide thickness, Source and Drain dimension, and many other parameters. This has worked well in the past but scaling is reaching its limits (8).

Some of the scaling problems arise due to material challenges, economic challenges, and device performance challengers (8). One example of material challenge is with scaling transistors channel length ever downward, the Gate insulating oxide is also scaled, which is usually composed of Silicon Dioxide, can no longer keep leakage current from tunneling from the channel to the Gate at a acceptable level. This leakage current is attributed to tunneling of channel electrons though the insulating oxide into the Gate using Fowler-Nordheim and direct tunneling. These quantum mechanical effects occur when sufficient electric field is
supplied by the Gate that channel electrons have a high probability of tunneling though the oxide layer (9). This is a major deleterious effect.

To overcome this high permittivity dielectrics are used, referred to ask high-k, which allows thicker insulating layers and reduces electron tunneling probability. This can be understood using the capacitance expression of the oxide layer between the Gate and channel expressed by 1.1 where $\epsilon_r$ is the relative permittivity, $\epsilon_o$ is the absolute permittivity, $A$ is the area of oxide layer and $t$ is the oxide thickness. As 1.1 shows that having a higher dielectric allows for having a thicker insulating layer while keeping the capacitance the same. With thicker oxide layer the Gate leakage current can be reduced (10).

$$C_{ox} = \frac{\epsilon_r \epsilon_o A}{t_{ox}} \quad (1.1)$$

Material challenge can also be expressed as a function of change in the increasing use of materials from the periodic table as shown in Fig. 1.1 (1). This advance has come at significant processing cost, when new materials are introduce they often require new machines and add to processing steps leading to complexity and adding cost.
With continued scaling and ever smaller channel lengths, new short channel effects are also seen such as subthreshold leakage current, Drain Induced Barrier Lowering (DIBL), punch-through and Hot Carrier effects (8–10). These effects in general occur due to contemporaneous scaling of the Gate supply voltage with channel length. While reduced supply voltage can save power they also reduce margin between supply voltage and transistor threshold voltage, the voltage where a transistor turns on. This threshold voltage is then susceptible to being altered by high Drain bias due to the short channel where the Drain electric field increases channel surface potential, this effect is called DIBL. Therefore the channel is no longer exclusively controlled by the Gate, as most long channel devices are, but also by the Drain. This is just one of the short channel effects among many.

One of the significant decisions the semiconductor industry has taken in recent years is to move to three dimensional transistors called Fin Field Effect Transistors (FinFET) (2). FinFET is often referred to as Trigate FET as the channel is and is seen in 1.2. Fig. 1.2 shows...
a comparison of a planer transistor and that of a FinFET, which has a three dimensional channel and Gate.

![Structural comparison between (a) planar MOSFET and (b) FinFET.](image)

The primary benefit of the FinFET over planer FET in reducing short channel effect is due to better control of the channel by the gate. The actual channel is no longer in the bulk substrate but rather is above it and the Gate metal is able to wrap around in three sides, this makes the Gate the dominant mechanism for controlling surface potential along the channel which helps in reducing the effects of DIBL (2).

A price is extracted by the FinFET for better short channel performance by having a
larger number of processing steps which increases cost. Fig. 1.2 makes clear in comparison to
the planer FET who much more complex the fabrication steps must be to produce a FinFET.
For low power applications and better short channel device performance, companies like intel
and AMD have been willing to pay the higher cost.

Multivalued logic is a branch of research which intends to use higher logic states than
current binary states of 0 and 1, by going to higher logic states such as ternary with 0,1,2
(00,01/10,11) and quaternary with 0,1,2,3 (00,01,10,11)(11). These devices have been ac-
tively explored and are being considered for future technology change when transistor scaling
is no longer able to deliver further computational power (12). future applications research
device have been fabricated and circuits applications demonstrated using a wide variety
of devices such as Resonant Tunneling Devices (RTD) (13), using Carbon Nanotube FET
(CNTFET) (14), Molecular Logic Devices using chemistry of molecules to provide logic
states(15), biological applications such as neuron MOSFET which mimics the behavior of
neurons (16), and many other concepts.

Quantum Dot Superlattice (QDSL) based devices has been fabricated with current-
voltage characteristics that demonstrate Multi-Valued Logic (MVL). Using QDSL layers
in the gate region (QDG) FETs, ternary and quaternary logic circuits have demonstrated
significant reduction in transistor count and lower power consumption. Floating gate Flash
memory QDSL exhibiting multi-bit storage has been demonstrated, doubling memory den-
sity. Although these transistors and memory devices have shown novel characteristics, they
require improvement in performance (current-voltage characteristics) for real world applica-
tions. In particular, QDG-FETs require wider intermediate saturation regions for low-noise
margin MVL logic and greater separation in voltage for multi-bit storage. One way to
overcome these constraints is to use more than two layers of QDSL. Assembly technique
is developed in this regard to assemble four layers of QDSL for Germanium and Silicon
Quantum Dots which is validated with Atomic Force Microscopy (AFM) data.

In terms of simulation, a Surface Potential based device model is developed, which in-
corporates the Density of States (DOS) calculation of QDSL derived from Kronig-Penney model, to simulate the behavior of MVL and memory device current-voltage characteristic. This is an improvement in accuracy from the current models used for circuit simulation but far less computationally intensive as quantum models used.
Chapter 2

SWS-FET Based Multi-Valued Logic Circuits

2.1 Spatially Wave-function Switched FET

Spatially Wave-function Switched FET is a unique multi channel transistor, unlike conventional transistor with one channel, the two channel and two Drain setup allows for unique device Mulit-Valued logic behavior. The MVL behavior is achieved by having two channels with their individual Drain current, $I_{d1}$ and $I_{d2}$ which also have their own regions of operation as a function of the Gate as well as one region when both are on. These three states of operation are the logic states that drives circuits based on this design.

MVL based devices will only be validated and realized when well established integrated circuits components are replicated, their performance verified, and benefits understood. With this intent, a design of Multiplexer and Demultiplex is under taken by using SWS-FET unique properties.
2.1.1 SWS-FET Device Characteristics

Multi-Valued logic devices have been created using Quantum Dots (QD) and Quantum Wells (QW), devices such as QDG-FET where the Quantum Dots are placed above the channel, QDC-FET where the QD are in the channel and are used to conduct the current. In addition a multi-channel QW based devices with two Drains has also been fabricated and demonstrated (17). The later device called a Spatially Wave-Function Switched FET (SWS-FET), has two channels that are separated by a barrier layer and also has two Drains. Figure 2.1 shows the device structure of a SWS-FET.

Figure 2.1: Quantum Well Based SWS-FET with two Drains for MVL logic applications.

SWS-FET has two channel which are individually connected to one of the two Drains. The two channels behave as QW due to their confinement by barrier layers, the channel is composed of InGaAs material and barrier layers is made of AlInAs. The device turns "ON" when a low voltage is applied to the Gate and a moderate voltage exists from the Source to each of the Drains. A current will start to flow between the Source through the bottom channel to Drain 1 while Drain 2 will not have any current. With sufficient increase in Gate voltage the lower channel current will over come the AlInAs barrier layer and current will be confined to the top channel and collected in Drain 2 with no current in Drain 1. Higher Gate voltage causes the current to switch from bottom to top channel due to electrostatic pull of the Gate and with higher electric field the electrons acquire sufficient energy to quantum tunnel into the top channel. There also exits a Gate voltage such that both channels are
conduct concurrently with both Drains collecting current.

Figure 2.2 shows quantum simulation of a SWS-FET where a 2D model is presented of a cross-section from the Gate through the two channels and into the substrate. The simulation results show that initially with low Gate voltage, the electron wave function or charges, are confined to the lower well as seen in 2.2a. A higher Gate voltage causes the electron wave function to tunnel into the top layer as seen in 2.2b.

2.3 presents the idealized current characteristics of the SWS-FET with logic states shown corresponding to current values in quaternary logic. 2.3a shows ideal $I_d-V_g$ characteristic where the Deep and Shallow Drain correspond to being connected to the bottom and top channel respectively. As stated before, the bottom channel conducts first and eventually this current will saturate, which is given a logic value of "01". With continued increase in Gate voltage, and above a certain threshold voltage $V_{th2}$, electrons will start to tunnel into the top channel causing both channels to conduct simultaneously, with a logic value of "11". Eventually with a high Gate voltage all current will be confined to the top channel and the saturated current is has a logic value of "10". In Fig. 2.3b shows an idealized model $I_d-V_d$ of the regions of Gate voltage and the corresponding channel current conduction and regions where only one or both channels will conduct.

Using the unique MVL characteristics of SWS-FET, a wide variety of circuits have been designed including Analog of Digital Converter (ADC) (18), logic circuits (19), and memories devices (20). In general by using SWS-FET with ternary or quaternary logic has been
Figure 2.3: Quantum Simulation of SWS-FET shows electron wave function confined to each of the channels.

demonstrated to provide significantly higher computational processing with fewer transistors which has leads to large power savings. One of the drawbacks of the SWS-FET has been the device area has to be kept large to accommodate two Drains which is an increase in device footprint. On benefit of larger device size is the relatively lower Gate leakage current compared to every shrinking conventional transistors.

2.1.2 BSIM Based SWS-FET model

To fully utilize SWS-FET devices, simulation of actual logic circuits are undertaken as cited above. This requires underlying device models, several approaches have been taken in this regard with one making the use of modified conventional transistors to behave as SWS-FET (21) and another using programming to mimic logic states (22). Both models do not capture the quantum behavior of the SWS-FET, as a such a model would prove too computationally intensive for circuit simulation purposes and nor do they accurately capture all of the resistances and capacitance values. These models and circuit simulation rather are meant for logic demonstration purposes.

To demonstrate the logic application of SWS-FET, a 2:1 and 4:2 Multiplexer and De-multiplexer design is presented (21). These designs makes use of Berkeley Short-Channel IGFET Model (BSIM) (23), which is a highly accurate and well known open-source compact
model of a conventional FET. The circuit simulations are carried out using ngSpice which is also a high fidelity open-source circuit simulator (24).

Two BSIM transistors are utilized where their Gates and Sources are held in common while each of their Drains, D1 and D2 separated and connected to their corresponding channels. Figure 2.4 shows a schematic of this design, the advantage of this type of model is that it has two individual channels which are then connected to each of their corresponding Drains, this is similar to how SWS-FET Drains are connected. Additionally the resistance of the channels and various capacitance values of the transistor are taken into account within the BSIM model. The current switching between the channels is controlled by setting the $V_{th}$ of each of the transistors to a different values, with D1 connected transistor having a $V_{th1} > V_{th2}$ compared to the D2 connected transistor. The difference in $V_{th}$ causes the transistors to turn ON for different Gate voltage, consistent with SWS-FET operation. Additional control is provided by using timing of the Source and Gate to insure the proper mimicking of SWS-FET current characteristics. This is necessary due to higher Gate voltages would keep both transistors ON.

![Figure 2.4: Two BSIM transistors are used to mimic twin Drain SWS-FET behavior using variable threshold voltage for each transistor and timing](image)

The disadvantages of this model should be noted, primarily this is done to demonstrate logic operations and the viability of a SWS-FET based circuits for lowering transistor
count, achieving low power consumption, and higher computational processing power. This model should not be used for accurate device characteristics for current output, resistance values, and capacitances. These parameters are inevitable different due to significant difference in device structures. A unique physics based compact model would have to be developed from first principals to incorporate the current switching between the channels as a function of Gate voltage and the related resistances and capacitances.

2.2 Multiplexer and Di-multiplexer

Multiplexers (Mux) and Di-multiplexers (DiMux) are conventional CMOS based devices which are widely used for routing signals that go into ports and switching them to proper channels (3). Applications ranging from switching signals in phones, to ADC, to routing data in memory buses. The advantage of using SWS-FET to develop Mux /DiMux is to reduce transistor count and thus lower power consumption.

The basis of operation of a conventional 2:1 Mux is that a single control signal $S$ regulates two inputs signals $A_0$ and $A_1$ to output port $Q$. This is shown in Fig. 2.5 where Fig. 2.5a shows the truth table of the inputs to corresponding output based on the logic value of the control signal. Figure 2.5b shows the conventional CMOS based circuit which consists of two AND and one OR gate with a total transistor count of 12. Similarly 2.6 shows a 4:1 Mux with two control ports, this design consists of two NOT, four three input AND, and one four input OR gate which roughly amounts to 38 transistors.
2.2.1 SWS-FET Based Multiplexer and Di-multiplexer Design

To show the use of SWS-FET can reduce transistor count in a Mux/DiMux circuit, an initial design is considered with only binary logic which facilitates comparison to current designs. A choice for ternary or quaternary logic could have been made which could reduce transistor count further but it would not allow back to back comparison to existing CMOS circuits.
Two approaches are chosen to develop SWS-FET based circuit, one is to use SWS-FET in conjunction with CMOS transistor and another is to design only use SWS-FET. These consideration are done with the following intent, if SWS-FET is used with existing CMOS technology then it could lead to ease of design for circuit development as most designers are vastly experienced in CMOS development. Additionally, SWS-FET only based Mux is developed with manufacture-ability in mind, as a mixed CMOS and SWS-FET would may contribute to significant cost increase from integrating Silicon with III-IV technology.

One of the strength of this design is that the DiMux is the same circuit as the Mux but with inputs and outputs reversed, therefore for a DiMux the input would be $Z$ and the outputs would be $A$ and $B$ with the control $S$ the same. This is not the case in a conventional CMOS based 2:1 Mux which cannot be reversed, for example a 2:1 conventional DeMux has 10 transistors compared 12 for the Mux.

![Diagram](image.png)

Figure 2.7: a) 2:1 Mux with an integrated CMOS and SWS-FET design and b) one with only an SWS-FET design. This SWS-FET only design can further reduce number of transistors and reduce manufacturing complexity from having to integrate CMOS and III-IV materials.

2.2.2 Design and Results for 2:1 Multiplexer and Di-multiplexer

Figure 2.7 is the schematic of a 2:1 Mux design. Figure 2.7a shows SWS-FET and CMOS integrated design with a pmos and nmos transistors connected to the $V_{DD}$ and the signal $S$
connect to their Gates. Depending on the value of $S$, either the nmos or pmos will activate causing two distinct Gate voltages to appear at the SWS-FET. The SWS-FET is graphically shown with a Source $Z$, the output of the Mux, and two Drains connected to their individual channels labeled $A$ and $B$ which are the inputs of the Mux. For the SWS-FET design only, Fig.2.7b shows the schematic of the circuit with an SWS-FET designed to operate on the same principles as a pmos transistor on top, with an input $A$ connected to the top channel and a regular SWS-FET with a top channel connected to input $B$ and both with common Gate and Source, with Source being the output $Z$.

The operation of the CMOS and SWS-FET integrated circuit is summarized here, initially with a logic 0 for $S$ the pmos is ON while nmos OFF, this causes the SWS-FET Gate to experience full $V_{DD}$ and causes the upper channel to be active, therefore the signal at $A$ will transfer to the output $Z$. With a logic 1 value for $S$, the nmos is ON and the SWS-FET gate experiences a voltage of $V_{DD} - V_{th}$, $V_{th}$ is that of the nmos transistor. This lower gate voltage activates the lower channel and the signal from $B$ passes to $Z$. Figure 2.8 shows simulation results done using ngSpice (24) where the output $Z$ correctly transfers the input results based on the $S$ value. This simulation result is exactly true for SWS-FET only design.
Figure 2.8: 2:1 simulation results for both CMOS and SWS-FET integrated circuit and SWS-FET only. The output $Z$ correspond correctly to the input $S$.

The DeMux circuit is the same as Mux circuits, as shown in Fig. 2.7 but operated with the inputs and outputs reversed. Figure 2.9 shows results for the DeMux circuit with output $Z$ correctly switching between the input $A$ and $B$ based on the control $S$ value.

Figure 2.9: 2:1 Demux simulation results for both circuits shown in Fig.2.7 with the inputs and outputs reversed
Design and Results for 4:1 Multiplexer / Di-multiplexer

In addition to a 2:1 Mux / DeMux circuit, a 4:1 Mux / DeMux is also created and simulated in ngSpice. A slightly different approach is taken with 4:1 Mux design where a four channels SWS-FET is used (20). Again two circuits are devised with one a combination of CMOS and SWS-FET and the other SWS-FET only. The four channel SWS-FET is used in conjunction with CMOS while two channel SWS-FET only circuit is used for the other design.

Figure 2.10a shows combination of CMOS and SWS-FET design with $S$ and $S_1$ being the selection bits and $A, B, C,$ and $D$ the outputs and $Z$ the input for a Mux design. Four channel SWS-FET behaves much like its two channel counterpart in that the channels farthest from the Gate has the lowest $V_{th4}$, therefore $V_{th4} > V_{th3} > V_{th2} > V_{th1}$. This property is taken advantage of with a combination of pmos and nmos to provide the SWS-FET Gate voltages corresponding to channel conduction. The circuit developed uses 6 CMOS transistor and 1 four channel SWS-FET.

![Figure 2.10: 4:1 Mux and DeMux circuits is designed with a) using a combination of CMOS and a four channel SWS-FET and b) an SWS-FET based circuit only.](image)

Figure 2.10b is an all two channel SWS-FET design and uses only 6 SWS-FET transistors. Both designs shown in Fig.2.10 are also DeMux circuits with their inputs and output reversed, which again is unlike their CMOS counterparts which have unique Mux and DeMux circuits. Figure 2.11 and 2.12 are simulation results for Mux and DeMux respectively,
both demonstrate the functionality of SWS-FET based design.

Figure 2.11: 4:1 Mux simulation for both combination of CMOS and SWS-FET and SWS-FET design only with output shown in $Z$.

Figure 2.12: 4:1 DeMux simulation results for both circuits shown in Fig.2.10 with the inputs being $Z$ and outputs $A, B, C, D$
2.3 Discussion

The Mux and DeMux circuits that have been constructed using SWS-FET demonstrate its potential for application in logic circuits. There are a few dramatic advantages with incorporating SWS-FETs logic and a few draw backs. One of the primary advantages that can be achieved is the significant reduction in transistor count which can lead to significant power savings. The 2:1 Mux, with combination of CMOS and SWS-FET circuit consists of 3 transistors, with SWS-FET counted as one, this is a 4X reduction in transistor count compared to conventional Mux which has 12 transistor. The SWS-FET design achieves a even higher 6X reduction. When 4:1 conventional Mux is compared to SWS-FET based designs the reduction is between 5X to 6X.
Chapter 3

Multi-Layered Quantum Dot Superlattice for Optimized Performance of Transistors and Memories

3.1 Quantum Dot Superlattice Novel Semiconductor Material

Quantum Dot Superlattice are defined as nanometer material individually confined in three dimension but when stacked together both horizontally and vertically will form a unique material with different properties than its bulk. This is due to the confinement being thin and allowing the electron wave function of a QD to merge with one next to it which gives rise to a unique band structure (25; 26). Lazarenkova et al does a particularly shows mathematically how three dimensional miniband formation occurs from a quantum mechanical calculation, this is an enhancement of the 1D analysis of the Kronig-Penny model (27).

Superlattices can be composed of a variety of materials, from III-V and II-VI materials
based QDs (28; 29), nanowires (30), and metal nanoparticle/polymer based material (31),
carbon nanotube (32) and many others.

We investigate uses of QDSL composed of Silicon and Germanium due to their ease of
fabrication and their self-assembly nature. Another significant advantage is that Silicon
and Germanium based QDSL are CMOS complaint and easily be integrated into current
fabrication process without requiring much change. This is not the case for III-V or II-VI
material or novel semiconductors like nanowires or carbon nanotubes.

Silicon and Germanium QDSL have been applied to transistors and memories and novel
characteristics have been shown. For application commercial products they must meet rig-
orous requirements which can be achieved by optimizing the QDSL stack.

3.1.1 QDSL Applied to Transistors

QDSL applied to transistors have shown MVL logic, where ternary and quaternary logic
states have been achieved (33–35). This is shown in Fig.3.1 where in 3.1a shows the structure
of a conventional FET and its resulting $I_D - V_g$ characteristics. The values for $I_D - V_g$
presented in 3.1 are all idealized and is intended to show contrast between devices and their
performance.

In addition when applied to memory devices, particularly in the Floating Gate of a Flash
memory device it has shown multi-bit memory storage per transistor.

With QDSL applied to the ”Floating Gate”, a region between the tunnel insulator and
Gate contact, the current has the ability to tunnel into the Floating Gate at voltages where
the current is able to tunnel through the insulator. The QDSL is then fills with charges as
a function of Gate voltage which causes the current in the channel to temporarily saturate.
When QDSL is completely filled and Gate voltage is at a higher value the current will
linearly increase and then fully saturate. FET’s made using these are called Quantum Dot
Gate (QDG-FET) and its idealized $I_D - V_g$ is shown 3.1b. The QDSL used in 3.1b are
made of formed from Si QDs, similar results can also be had if only Ge QDs are used.
Figure 3.1: Transistors with QDSL applied show significantly different $I_D - V_g$ which is ideal for MVL application.

The unique $I_D - V_g$ of this device represents ternary logic where there is an OFF state (00), an intermediate state (01 or 10) and a ON state (11). Figure 3.2a shows how current tunnels into the QDSL layer and subsequently leaks to the Gate as a function of Gate voltage while 3.2b gives the Drain current profile. The Drain current temporarily saturates to form the intermediate state when charges are accumulating in the QDSL layer causing the surface potential in the channel to not change even while Gate voltage keeps increasing. The temporary saturation of the surface potential along the channel as a function of QDSL charge can be explained from a charge conservation and electrostatics point of view.
In Fig.3.1c, the Floating Gate consists of Si and Ge QDs, with Ge assembling on bottom and Si on top, this is referred to as a "Mixed" dot configuration. The $I_D - V_g$ of the device has an additional intermediate current saturation region, initially current tunnels to the bottom Ge layer causing the first intermediate state, with increasing voltage the current is able to overcome the bandgap difference between the Ge and Si QDs and are stored in Si QDSL leading to second intermediate state. With enough Gate voltage the current will leak into the Gate causing saturation. Like the single layer QDSL based QDG-FET, mixed dot configuration has quaternary logic values with ON, intermediate 1, intermediate 2, and OFF state representing logic values of 00, 01, 10, and 11 respectively.

Figure 3.3a and b shows fabricated device structure and $I_D - V_g$ characteristics of a four state transistor (34). These results demonstrate the viability of these device to produce unique ternary and quaternary results. The $I_D - V_g$ clearly show intermediate states, $i_1$ and $i_2$, which are saturation regions in the Drain current.

Although these are novel results, these devices need further refinement and optimization for them to be truly applicable in circuits, the primary concern is that these intermediate states, explicitly $i_2$, is not wide enough and not separated enough from the ON state to meet noise margin requirements between states.
3.1.2 QDSL Applied to Flash Memory

QDG-FET based Flash memory devices have also shown promising results (34; 36), where QDSL are again used in the Floating Gate region but capped with an insulating layer from the Gate. Figure 3.4a shows the conventional Flash memory structure with a charge storage layer juxtaposed between insulating layers. This type of device is often made with Silicon Oxide Nitride Oxide Silicon (SONOS) with the charge storage element being Silicon Nitride ($Si_3N_4$) (37).

General Flash memory $I_D - V_g$ operation is shown in 3.4a where an initial read cycle returns the typical transistor current profile. A write pulse is then initiated which applies a high enough voltage at the Gate and for a certain duration which allows charges to tunnel from the channel into the Floating Gate layer but is prevented from leaking into the Gate by the Gate insulator. This stored charges causes a shift in the $V_{th}$, when another read cycle is initiated a shift in the $I_D$ is seen due. From the difference in current value as a function of voltage difference from the initial read to post-write read phase, a single bit of memory stored is identified.

Figure 3.4b shows the use of Mixed dot QDSL in the Floating Gate region. This leads to
multi-bit storage because of the ability of the Si and Ge QDSL layers to individually store charges, much the way that a QDG-FET "Mixed" dot configuration did for MVL circuits. An initial read phase returns the typical transistor characteristics, a first write with enough voltage is applied to only bring charges to the bottom Ge layer and produces a large $V_{th1}$. The post-first write read phase will show a typical Flash memory characteristics as seen in 3.4a, with a second write pulse applied with a higher voltage allows the charges to tunnel into the Si QD top layer, producing a second $V_{th2}$. This causes the subsequent read phase to show a backwards shift in the $I_D - V_g$ characteristic as $V_{th1} > V_{th2}$. The $I_D - V_g$ characteristic of the Mixed QDG-NVM from the first write to the second write is shown in 3.4b and is indicated by arrows.

Figure 3.4: QDSL applied to Flash memory can optimize charge retention and show multi-bit storage

Replacing the Floating Gate layer with nano-crystal has been demonstrated by Tiwari et al. (38–40). The primary advantages of using nano-crystals is that they are charge trapping, are self-contained which helps in charge retention, and can often be deposited in
controlled and uniform manner. The SONOS use of \((Si_3N_4)\) relies on the natural nano-crystals formation during low temperature Chemical Vapor Deposition (CVD), in essence these nano-crystals are regions of crystalline materials surrounded by its amorphous phase. Si and Ge based QDSL are also classified as nano-crystal material for memory applications. The primary difference between SONOS application and QDSL is in cost, uniformity of deposition, and memory retention.

Figure 3.5: Fabricated device structures in a) produces the \(I_D-V_g\) results shown in b). Multi-bit storage in the multiple \(V_{th}\) shift seen.

Figure 3.5 a and b show the fabricated device and its resulting NVM characteristic, respectively. Figure 3.5b shows two bit of memory storage in one device with arrows indicating that first bit creates the wider \(V_{th}\) shift with the second bit contracting it. Again these are novel results, particularly as they double the memory density while only slightly increasing the vertical area.

QDG-NVM, like QDG-FET, require refinement and optimization for it to find application in real world condition. The separation between \(I_D-V_g\) of the first and second stored bit is far too small for it to meet noise margin requirements. The noise margin establishes the minimum \(V_{th}\) shift between the two storage states for them to be distinguishable.
3.2 QDSL Charge Storage Capability

Both QDG-FET and QDG-FET show very novel and promising results for MVL logic applications and higher memory density, respectively, yet both suffer from having suboptimal charge storage capability. With inadequate level of charge storage the QDG-FET shows intermediate states which are not as distinct as they need to be for device applications. This can be visualized by comparing the idealized performance in 3.1c and the actual device performance in 3.3b. Similarly, by comparing 3.4b and 3.5b shows the contrast in $I_D - V_g$ performance desired and actual results.

One of the ways to get desired performance must be to understand the charge storage capability of Si and Ge QDSL individually. This can be done by fabricating a Metal Oxide Semiconductor Capacitor (MOSCAP) of the same dimensions for Si and Ge and comparing their charge storage capability. A cross-section of a MOSCAP fabricated is shown in 3.6a where Si or Ge QDSL sits on a p-type substrate with their individual cladding as the only tunnel insulator and they are capped with 35 nm of Hafnium / Aluminum Oxide stack to prevent charges leaking into the Gate. The MOSCAP has a Aluminum Gate and back contact which can be used for Capacitance-Voltag (C-V) measurements. In order to definitively show that the Gate Insulating layer does not act as a charge trapping layer a MOSCAP with exact dimensions and materials but with the QDSL layer missing is fabricated. This is necessary as of then Silicon Nitride has been used as a Gate Insulator and it is a well known material to form charge trapping nano-crystals. By using a Hafnium / Aluminum Oxide stack which is less prone to formation of these nano-crystals, it can be shown that all charges are trapped by the QDSL layer.
Figure 3.6: a) The structure of MOSCAP consists of the QDSL layer with only the cladding acting as tunnel insulator and $\text{HfO}_2/\text{Al}_2\text{O}_3$ stack acting as a Gate Insulator. b) Demonstrates that control MOSCAP has virtually no charge storage capability.

The C-V measurements that are carried out for charge trapping study are all high-frequency measurements, those at 1MHz, and the C-V is a sweep from negative to positive voltage and then this is reversed. This method of sweeping the voltage allows charges to tunnel into the QDSL layer and cause a $V_{th}$ which shows up as hysteresis. Figure 3.6b shows virtually no hysteresis which indicates that the Gate Insulating layer is not trapping any charges. In Fig. 3.7 shows results for MOSCAP with Si and Ge QDSL, and clearly hysteresis is seen in both cases. This indicates that these QDSL are charge trapping, with Si having roughly 5 times the capability of Ge. This is somewhat expected as Ge QD are half the size of Si QDs, when volume of the inner core is taken into account Si QDs are 8X larger than Ge. The charge storage capability does not directly corresponds only to size but also Density of States (DOS) which will be discussed in later sections.
Figure 3.7: Si and Ge MOSCAP show hysteresis which indicates charge storage with Si QDSL storing significantly more, this gives designers a chance to chose QDs used based on charge storage needs.

Several broad conclusions that can be drawn from these MOSCAP C-V measurements is that charge trapping capabilities between Si and Ge are significantly different with Ge having 5X less charge storage capability. This information allows NVM designers to chose which QDSL to use based on the desired charge storage capability. This study also demonstrates that the cladding materials of the QDSL are enough for charge storage, the MOSCAP structure in different from a NVM as it lacks the channel insulator. Therefore the QDSL are self contained charge trapping devices, this can be particularly useful for applications with lower dimensions where QDSL can be assembled without any need for channel insulator, saving processing steps, time, and reducing cost. Additionally, the MOSCAP control sample validates NVM devices with QDSL layers and Hafnium / Aluminum Oxide stack that they are the primary charge trapping mechanism in NVM. Ge QDSL also shows a much higher initial capacitance value compared to the Si, which indicates that the greater dielectric coefficients of the Ge and Si can be utilized as a design tool for further analysis.
Figure 3.8: Low Frequency C-V shows charges tunneling into QDSL layer, causing capacitance to go down temporarily

In Fig. 3.8 shows a low frequency (100 KHz) C-V measurement taken on the Ge QDSL MOSCAP, this shows charges tunneling into the QDSL layer when Gate voltage reaches a level where it is able to overcome the cladding of the QD and gets trapped in the core. This causes the capacitance to drop, with higher gate voltage the QDs are fully charged and the expected increase in capacitance at low frequency is found. It should be noted that the charges can be seen tunneling at low frequency but not at high frequency due to the inversion layer never forming which keeps the capacitance constant.

For QDG based transistor and memory to become realizeable, their intermediate states and greater $V_{th}$ for multi-bit storage has to be increased. This can be done by including more than two layers currently used, this would require substantial change in the way that current QDSL are assembled.
3.3 Quantum Dot Supper-Lattice Characteristics

QDSL made from Silicon (Si) and Germanium (Ge) are used where the individual Quantum Dots (QD) are of the diameter of 6 nm to 3 nm respectively. These QDs have unique properties such as higher bandgap compared to their bulk, this is due to confinement where the QDs are composed of a inner intrinsic core and are surrounded by a cladding layer. Figure 3.9a and b shows Si and Ge QDs cross-section

![Figure 3.9: Self-Assembled QD will only assemble over p-type semiconductor](image)

Quantum Dot Supper Lattice (QDSL) structures have been studied for decades, with most being of the quantum well type. The QDSL structure studied and modeled here is the based on Self-Assembled Quantum Dots of Silicon and Germanium, which assemble in 2 layers. This type of QDSL is unique in that it does not require CVD equipment to be used in its formation but rather are fabricated by low cost equipment and assembles on a wafer when it is submerged in a QD containing solution.
Figure 3.10: a) Si QDSL will only assemble on p-type wafers in 2 layers and b) it is desired to get to 4 layer of QDSL assembly of only Si or Ge or a combination of the two.

Figure 3.10a shows how QDSL assembled on a p region of a wafer, the QDSL will only assemble on the p region due to the cladding chemistry which is electro-statically attracted to positive regions. The
Figure 3.11: a) An AFM surface profile of a line sample with Germanium QDSL assembled, dark band represent where QDs did not assemble as seen in (b) where an average surface profile cut is shows, this shows Germanium QD sizes to be 4 nm. A three dimension surface profile of Silicon QDSL is seen in (c) with some sample defects and AFM data collection artifacts but an average surface profile cut shows in (d) that Silicon QDs to be 6 nm.

The QD’s will only assemble over p-type semiconductor due to an electrostatic attraction resulting from a negative charge found in the cladding layer from the $SiO_x$. This property is found in both the Silicon and Germanium Dots. Furthermore, single QDs are attracted to other QD in such a way that they will only Self-Assemble in two layers at a time. This electrostatic attraction is taken advantage of to assemble the QD over p-type channel regions, which makes it site specific, which is why this method is most often refereed to as Site Specific Self-Assembly.

The unique characteristics of QDSL arises due to the confinement of intrinsic Silicon and Germanium, with bandgap equal to 1.12 ev and 0.64 ev respectively, by Silicon Oxide and Germanium Oxide cladding layer with bandgap of 8.9 ev and 6.0 ev respectively. If the bandgap profile of one of the layer of Silicon QDSL is looked it, in one direction, it would
resemble the bandgap of a stack of Quantum Wells, this is shown in figure 7.

Figure 3.12: Bandgap profile of one layer of Silicon QDSL resembles a Quantum Well band structure.

An individual quantum well exhibit discreet states due to the electron wave function having discreet solutions to the Schroedinger wave equation. When many quantum wells are stacked together, if their barriers are thin enough then the wave function will extend into another quantum well, thus the wave functions of each of the discreet levels merge and form minibands. These minibands constitute energy levels where the electrons can easily travel through the quantum well.

These minibands, in effect, create a set of new conduction and valence band structure. To ascertain the profile of conduction and valence bands, the Kronig-Penny model has to be solved for a QDSL structure in three dimension, or the Brillouin zone [Ref]. This reveals minibands have a certain width to them, which dictates the carrying capacity for electrons in each of the minibands. This was done for both Silicon and Germanium QDSL and the three most relevant minibands are shown.
The width of the minibands can dictate the density of states of each miniband, this gives the number of states in which electrons can exists. The density of states can be found by using equation 3.1, where the $E_c$ represents bottom of the miniband and $E$ a particular energy level or top of the miniband.

$$g_{mb}(E) = \frac{1}{2\pi^2} \left( \frac{2m_{eff}}{\hbar^2} \right)^{3/2} \sqrt{E - E_c}$$ \hspace{1cm} (3.1)

The density of states formula used here is for that of bulk material, there are density of states function for both Quantum Wells and QDs but our QD’s in QDSL are not true zero dimensional objects. There is an ongoing debate about which density of states function should be used, so far treating the minibands with bulk density of states functions is done. The density of states are filled when an applied energy is applied which causes the electron to transit to the minibands. This can be calculated by applying from an arbitrary starting point a set of energy which increases the fermi level,, equation 3.2, until bands are filled. The total number of electrons stored in the mini bands can then be calculated as a function of the fermi level and the density of states of the bands, as seen in equation 3.3.
\[ f(E) = \frac{1}{e^{\frac{E-E_c}{kT}} + 1} \quad (3.2) \]

\[ Q_{FG} = \int_{w_1}^{w_2} f(E) g_{w_1}(E) + \int_{w_2}^{w_3} f(E) g_{w_2}(E) + \int_{w_3}^{w_4} f(E) g_{w_3}(E) \quad (3.3) \]

Figure 3.14 shows results of simulated calculation of applying a theoretical gate voltage, which increases the fermi level, until charges fill minibands completely. This is only carried out for the lowest three minibands, Silicon and Germanium QDSL have many more. Summation of charges across the minibands, in essence executing equation 3.3, gives total charge as a function of gate voltage for Silicon and Germanium QDSL, shown in Fig. 3.15. With total charge storage in the minibands known as a function of applied voltage, this would theoretically allow intermediate state widths to be calculated for MVL FET applications and threshold shifts for memory applications, both parameters critical to their operations.
Figure 3.15: Total charge in minibands calculated for Silicon and Germanium QDSL for an applied voltage

The miniband formation alters the intrinsic bandgaps of Silicon and Germanium, where the energy required for electron transition from the highest valence minband to the lowest conduction band increases. In essence, this is bandgap engineering, by making the QD’s of the QDSL smaller the quantization can be enhanced producing a larger badgap. This larger bandgap is verified using photoluminescence (PL) measurement where a poly-silicon wafer is has two layers of Germanium QDSL assembled on it and then is excited by photons with wavelength of 500 nm, this causes excited electrons to jump to the altered conduction band, when they relax back down to altered valence band a photon is emitted. This emitted photon is then measured to reveal the bandgap. Figure 3.16 shows the results of PL measurement where a Germanium QDSL peak is seen at roughly 1280 nm wavelength which corresponds to 0.97 eV bandgap. Similar measurement and theoretical calculations have shown Silicon QDSL to have a bandgap of 1.24 eV. The altered bandgaps when compared to their bulk material are 10 and 45 percent higher.
One of the exciting aspects of a general Superlattice is that it is a metamaterial, in that it is not generally found in nature but is man made with altered properties compared to its bulk. The cause of the altered properties have been discussed which relates to altering of band structures and new bandgaps emerging. For a QDSL these properties are dependent on the size of individual QDs and the ratio of core to cladding. If the cladding is too thin then they and cores large then they will more resemble their bulk counter parts, if the cladding is too thick then mini bands may not form and quantum tunneling would be too prohibitive. Getting the core to cladding ratio correct can lead to miniband formation and reducing the QD total size can increase bandgap.

In addition to a altered band structure and bandgap, QDSL metamaterials also exhibit an altered permittivity. This arises due to the nanometer nature of the material and the corresponding photon whose wavelengths are much larger. For example, a Silicon QD has a core the size of 4 nm with 1 nm of Silicon Oxide, when PL measurement is is taken with wavelengths as low as 200 nm, relative to the QD size the wavelength is 40 times larger.

Figure 3.16: Photoluminescence measurement shows altered bandgap of Germanium QDSL at 0.97 eV compared to Bulk value of 1.24 eV.
and cannot distinguish between the intrinsic Silicon core and Silicon Oxide but rather sees a uniform compound. This new compound will have an effective permittivity that is a function of the permittivities of the two compounds and their volume fractions, this is given by the equation below.

\[
\frac{\epsilon_{\text{eff}} - \epsilon_{\text{clad}}}{\epsilon_{\text{eff}} + 2\epsilon_{\text{clad}}} = f \frac{\epsilon_{\text{core}} - \epsilon_{\text{clad}}}{\epsilon_{\text{core}} + 2\epsilon_{\text{clad}}}
\]

The above equation is called the Clausius-Mossotti formula (41) which is calibrated for spherical objects and it gives a reliable estimation of the effective permittivity based on the core and cladding permittivities and their volume fraction \(f\). Using this to calculate Silicon QDSL at 6 nm QD with core size of 4 nm and 1 nm of cladding of Silicon Oxide, \((\epsilon_{\text{core}} = 11.68, \epsilon_{\text{clad}} = 3.98)\), this gives effective permittivity at 7.4. Similarly for Germanium QDSL with 4 nm QD with core size 3 nm, \((\epsilon_{\text{core}} = 6, \epsilon_{\text{clad}} = 16.2)\), and gives the effective permittivity as 9.24. This shows that depending on the core and cladding ratio, the permittivity can be controlled.

### 3.3.1 QDSL Fabrication

The chemistry behind formation of the cladding layer and its negative charge is briefly described here for Silicon QDs only but Germanium QDs follow similar procedure, more detailed source can be found here (5; 42) Once QD solution is made, it is kept in the sonicator prior to use, the sonication process activates oxygen, the benzoyl and siloxy radicals react with ethanol forming benzoylethylester, benzoic acid, tetraethoxysilane (TEOS), and H2O, which is shown in Fig. 3.17. The benzoylethylester correlates to the release of H2O which causes slow oxidation. Further functionalization of the outer layer of QD is done by generated TEOS with Si-OEt groups in portions of silonols, the Si-EtO group remains unhydrolyzed.
due to the consumption of H2O during oxidation (43).

![Diagram](image.png)

**Figure 3.17: Sonication assisted oxidation of quantum dots forming the cladding layer.** (5)

There is a strong dependency of PH of the QD solution to its suspension and eventual assembly on the surface a substrate. Experimentation has shown that when QD solution is kept between PH 7.8 to 6.7, the suspended QD remain stable and do not amalgamate. The negative charge attributed to the QD’s are due to silonal ionization. The QD will amalgamate and precipitate for PH between 6.7 to 4.9 due to charge neutralization of Si/SiOx colloids. Below PH 4.9 the QDs in solution will once again suspend due to partial protonation and now will precipitate as a uniform layer on a p-doped substrate. Fig. 3.18 summarizes the QD suspension and precipitation as a function of PH. The negative charge of the QDs attract them to the positive or p-doped part of any substrate only due to electrostatics. Once two layers are assembled, forming a QDSL, the top layers prevent further formation due to electrostatic repulsion between the negatively charged QDSL layers and the negatively charged suspended QDs.
Due to the redundant nature of describing QDSL fabrication, the above section is taken from a forthcoming journal article in Journal of Electronic Materials titled 'An Investigation of Quantum Dot Super Lattice Use in Nonvolatile Memory and Transistors.' and written by this author.

Following the above procedure the QDSL solution is ready for use in self-assembly. When QDSL solution is applied to a wafer, it is first dipped in dionized water, then in methanol upon which it is dried using nitrogen and the wafer is then immersed in the solution. The wafer is left in the solution for anywhere from 3 to 6 minutes during which time the QD assembles on to the p-type regions. The wafer is then taken out and dipped in methanol before dried with Nitrogen. The solution is then annealed at 700 °C for Silicon, 250 °C for Germanium, and 325 °C for a combination for Germanium and Silicon. This annealing steps helps in adhesion of the sample to the wafer.

This procedure will self-assemble two layers of QDSL on a wafer, for both Silicon and
Germanium, and a combination of one layer each of Germanium and Silicon. AFM has been used to verify the two layer formation of Silicon and Germanium is shown in 3.11.

3.3.2 QDSL Four Layer Fabrication

There maybe advantages for using four layers of QDSL instead of the current two for better device performance to, for engineering new bandgap materials, and optical applications. Therefore a technique has been devise with a combination of intuition and trial and error which are then verified using AFM.

It has been found that four layer of Germanium can be self-assembled if the two self-assembly procedures are conducted including the annealing steps. A modification is added where between the annealing steps the wafer is submerged in deionized water for a period of 3 to 6 minutes. This step is crucial as the deionized water is presumed to help in developing OH bonds on the cladding layers of already deposited QDs. The further development of the OH bond then helps in binding more QDs on the the existing deposited QD, which are also electrostatically assisted by the p-type attraction. It is also critical to use fresh QDSL solution for each self-assembly and not reuse solution.

This procedure works well with Germanium and two layers of Germanium assembled on two layer of Silicon. It is presumed that this will also work well with making four layers of Silicon, and alternation layers of Silicon and Germanium.

In addition to developing a methodology for building four layers, a single layer of QDSL has been demonstrated. This has been done by reusing the QDSL solution, the first step is to use the solution to deposit two layers of QDSL on a wafer. This depletes the available number of QD in the solution and therefore the density of QD in solution goes down. A second self-assembly is conducted with this depleted solution on a new identical wafer, the combination of low QD density and large p-type region with QD wanting to be at the lowest electrostatic potential allows the formation of a single layer of QDSL. Single QDSL layer has been verified with AFM.
It is suggested that to form a single layer of QDSL, one needs to only control the concentration of QDs in the solution. This can be done by using only a quarter of the QDSL solution normally used for two layer deposition and rest filled with ethanol to get to a full solution. This new solution should have quarter of the QD density as the original solution. This process would obviously have to be refined with repeated self-assembly and verification by AFM but the possibility of single layer QDSL has been demonstrated.

### 3.4 QDSL Optimization

Current fabricated QDSL based devices such as MVL four state QDG-FET seen in Fig. 3.3a while novel does not have ideal ID-VG characteristics for practical circuit applications. This can be seen when comparing the device performance to idealized characteristics shown in Fig. 3.19. The ideal case has distinct and clearly separated logic states while actual device has an intermediate state, $i_2$ that is not well separated from the full saturation state.

![Figure 3.19: a) Fabricated QDG-FET results have non idealized intermediate states b) compared to idealized states which are more distinct.](image)

When circuits are designed, transistors performance must meet clear metrics, these metrics are there to ensure that device to device variations, non optimal operating conditions, and device degradation will not adversely affect the circuit for duration of its life cycle. One of the critical metrics is that logic states must be distinct and well separated which the ID-VG curve in Fig. 3.19a would not meet but what is desired is shown in 3.19b.
There is also a similar case that can be made with the QDG-NVM device in Fig. 3.5. In Fig. 3.20a shows real two bit storage capability for a mixed dot QDG-NVM, when this is compared to idealized case in 3.20b where the ID current shows distinct Write states. The non distinct write states in 3.20a presents a challenge, the rate of error in detecting the wrong stored Write state would be too high for practical application. Device performance therefore would have to resemble the idealized cases.

As the arrows indicate in Fig 3.19 and 3.20, while fabricated results for QDG-FET and QDG-NVM are novel, they are need to be optimized to meet metrics for circuit application, in other words they must behave more like their idealized cases.

It is proposed that this can be achieved by increasing the stored charge states in the QDSL layers by going from two layers to four layers, in the mixed dot configuration. This would entail devising a process by which two layer of Silicon or Germanium QDSL could be grown on top of the other. Figure 3.21 shows what a four layer QDG-FET would look like in 3.21a and what a QDG-NVB in 3.21b.
By going to four layers and using a combination of both Silicon and Germanium, distinct states can be had. This is because with two layers of QDs would store more charges compared to one layer, leading to bigger intermediate states and larger threshold shifts. Additionally using both types of QD takes advantage of their inherent band gap difference with Silicon QDs at 1.24ev and Germanium at 0.97ev, this should provide a distinct step between intermediate steps and would add to the threshold shift. The challenge to making four layers is modify the existing process to assemble four layers as current techniques can only produce two layers.

Figure 3.22: QDSL made with a bottom two layers of Silicon and two layers of Germanium on top.
Using the fabrication modification techniques described in QDSL Four Layer Fabrication, line samples were used to assemble two layers of Germanium on top of two bottom layer of Silicon. The critical steps that allowed this to happen was to perform two separate self-assembles on one wafer including annealing but leaving it in deionized wafer after the first anneal and before the second self-assembly. This allowed the formation of OH bonds in the cladding layer which allowed further QDs to bond on top of the two layers. Figure 3.23a shows the QDSL stack and the expected heights while 3.23b and c shows a 3D AFM views of the QDSL dispersed on the wafer and the cross-section cut showing how height of the stack. As expected, the height of the stack should be 18 nm and the cross-sections shows this but also shows some non uniformity.

Figure 3.23a shows Germanium assembled on top of Silicon, this procedure is also expected to produce Silicon on top of Germanium and alternating Germanium/Silicon stacks. This assumption is made on the basis that the cladding chemistry remaining same for both Silicon and Germanium in that OH bonds play a central role.

Figure 3.23: a) QDSL made from four layer of Gemanium with b) showing 3D AFM profile and c) showing a cross-section profile of heights along the wafer. Four layers can lead to increased charge storage capability.

Figure 3.24 shows AFM results for four layers of Germanium assembled on a line sample,
the expected value seen in 3.24a is verified in height cross-section in 3.24c. This stack can be applied to devices or used for optical applications. One of the advantages of going from two layers to four layers is that two layer QDSL are superlattices in only two dimensions with height being a main restriction. With four layers the height allows for more stable bandgap in superlattice structure. Additionally single QDSL four layer stack can be useful in fabricating waveguides, lasers, and solar cells (44; 45). Four layer QDSL stack was only done for Germanium but this can also be extended to Silicon and for the same reasons as supplied for four layer for mixed QDSL configuration.

Figure 3.24: a) Single layer of Silicon QDSL formed on a wafer with b) showing a noisy AFM image but cross-section cut c) is taken where in a region with very little noise. The single layer of QDSL is formation is predicated on significantly lower the QD density of the solution.

While investigating and developing process of growing four layer stack of QDSL, a single QDSL layer growth was found. Figure 3.24 shows AFM results for one layer of Silicon QDSL, although data collected may appear noisy, the cross-section cut was taken from a less noisy area. The technique to single layer formation is to control the QDSL solution QD density. By reusing a solution already used to make two layer QDSL, the QD density in the solution is very low. The solution can then be used with a line sample to form a single layer of QDSL. Although applications of single layer is yet undefined there maybe future applications where single layer QDSL is vital.
This process of single layer QDSL assembly can be more formalized by starting with a quarter of the original QDSL solution and topping it off with ethanol to make up the original volume. Ethanol does not really effect the QDs but is used as a suspension agent. Further experimentation and refinement with QD density would produce a process with high fidelity. This would then have to be verified using AFM.

3.5 Discussion

Having better device performance characteristics are critical for QDSL based transistor and nonvolatile memory being implemented. The two layer QDSL that were used in prior devices provided novel and unique performance and showed that QDSL can be a solution to the scaling challenge that computing semiconductor industry is facing. For adaption of a technology by semiconductor industry it must be robust where and must meet metrics to allow it to operate correctly at worst condition situations.

Two layer QDSL showed some drawback due to having intermediate states that were not distinct enough and two bit nonvolatile memory having threshold shifts not big enough. This can be solved by going to larger QDSL with two distinct QDSL layer, this would provide more charge storage capability and a distinct bandgap difference which contributes to larger intermediate steps and greater threshold shifts. A clearly processing methodology was developed to get to 4 layers and verified using AFM results. This can now be implemented in transistor and memory devices verify the more robust results expected.

Another unique aspect shown by QDSL is the unique bandgap and permittivity as a function of the size of the individual QDs. These properties are useful, for example, in the optical application (46). Although this is has not been thoroughly explored, there remain a large possibility as some PL measurements verified the elevated bandgap. In addition it is presented below a transmission measurement of a poly silicon sample with and without two layer Germanium QDSL on it which shows a shift, as seen in 3.25. Although a true
absorption measurement of Germanium QDSL would have to be made for conclusive analysis, the transmission shift is expected due to the altered bandgap and permittivity. What this shows is the possibility of unique applications of QDSL in the optical regime.

Figure 3.25: Transmission measurements of poly silicon wafer with and without 2 layer of Germanium QDSL. The shift in transmission shows the effect of Germanium QDSL. Additionally previously collected Silicon absorption spectra is shown.
Chapter 4

Surface Potential Model of QDG-FET

4.1 Surface Potential Based Quantum Dot Model

Surface potential based compact model offers one of the more rigorously accurate physics based approach to compact modeling while requiring substantially less computation time compared to a quantum model. The surface potential based compact model was first proposed by Poa-Sah in 1965 (47) and have become a standard by which other compact models compare their results to for accuracy. The benefits of the Surface potential model is that it is an all region model, incorporating both linear and saturation behavior of a MOSFET. This is typically where most model differ, they break the linear and saturation regions into two distinct equations and solve them separately but this usually requires a fitting parameters to achieve the accurate Id-Vg characteristics but requires far less computation.

The surface potential model is a very suitable starting point for modeling MVL based devices due to its physics based approach and high accuracy when compared to experimental results. Although this approach can be computationally intensive, further optimization can lead to quicker computation time but as a first approach to modeling it is imperative to capture the device behavior before computational time is considered [revisit this paragraph]

The Surface potential model originates from solving the Poisson’s equation and conser-
vation of potential and charge for a MOSFET. As suggested by the name of the model, the potential of the interface between the tunnel layer and the semiconductor is a critical parameter which governs the current-voltage characteristic of the device.

Our aim here is to derive the Surface Potential model, starting with that of a conventional MOSFET and then making critical changes which will turn it into a model for a QDGFET. The below derivations for a conventional mosfet are well outlined in (48; 49). We first start with conservation of potential in a MOSFET, given by 4.1, where $V_{gb}$, $\psi_{ox}$, $\psi_s$, $\phi_{MS}$, is the applied gate voltage, tunnel oxide potential, surface potential, and metal-semiconductor work function respectively. The back contact of a MOSFET is assumed to be connected to ground and therefore the applied gate potential must be conserved by potential drops in overcoming the metal-work function, the drop in the tunnel oxide layer, and in the drop channel.
Equation 4.2 gives conservation found in gate, tunnel layer, and channel charge given by the variables $Q_g$, $Q_o$, $Q_c$ respectively. The gate charge occurs due to application of a gate potential, the tunnel oxide layer charge is often due to trapped charges present in the oxide layer, and channel charge occurs due to the applied gate voltage causes the channel into depletion and inversion. Note that tunnel layer is used to identify what in conventional MOSFET is often called gate dielectric, tunnel layer phrasing is chosen so to remain consistent when describing QDGFET.

$$V_{gb} = \psi_{ox} + \psi_s + \phi_{MS}$$  \hspace{1cm} (4.1)

$$Q_g + Q_o + Q_c = 0$$  \hspace{1cm} (4.2)

The channel charge, $Q_c$, can be divided into two components, the inversion layer charge, $Q_I$, which is only a sheet layer of charge at the interface of tunnel oxide and semiconductor layer and the charge in the semiconductor bulk, $Q_B$, which is primarily concentrated in the depletion layer. This is shown in 4.3 and 4.2 is restated using this relationship to give 4.4.

$$Q_c = Q_I + Q_B$$  \hspace{1cm} (4.3)

$$Q_g + Q_o + Q_I + Q_B = 0$$  \hspace{1cm} (4.4)
Using the above conservation of potential and charge, we can define the tunnel oxide potential, $\psi_{ox}$, and the metal work function $\phi_{MS}$, as a function of charges, $Q_g$, $Q_o$, flatband voltage, $V_{FB}$, and tunnel oxide capacitance $C_{ox}$, $\epsilon_r$, $\epsilon_o$, $t_{ox}$, are the dielectric of the tunnel oxide, the permittivity constant, and thickness of oxide layer, respectively.

$$\psi_{ox} = \frac{Q_g}{C_{ox}}$$  \hspace{1cm} (4.5)

$$\phi_{MS} = V_{FB} - \frac{Q_o}{C_{ox}}$$  \hspace{1cm} (4.6)

$$C_{ox} = \frac{\epsilon_{ox}\epsilon_o}{t_{ox}}$$  \hspace{1cm} (4.7)

$$Q_B = -\gamma C_{ox}\sqrt{\psi_s}$$  \hspace{1cm} (4.8)

Figure 4.3 shows an NMOS with the relevant charges as stated in /reffullchargeconserve. The charge in the inversion layer, $Q_I$, is significantly greater than charge in the bulk, $Q_B$, this has the effect that often models will consider the charges in the inversion layer to be only at the interface of the tunnel oxide layer and the semiconductor layer. Modeling an infinitely small charge at the interface turns out to be a very computationally efficient and simplifies many of the equations.
Having accounted for $C_{ox}$ and $\phi_{MS}$, we can also find $\psi_s$, by redefining 4.1 for surface potential using 4.5 and 4.6, which gives 4.9.

$$\psi_s = V_{GB} + \frac{Q_{g} + Q_{c}}{C_{ox}} - \phi_{MS} \quad (9) \quad (4.9)$$

Poisson’s equation 4.10, is a summation of the electrons concentration, $n(x, y)$, hole concentration, $\rho(x, y)$, and doping of the channel, $N_A$ along the channel, from Source to Drain, and into the bulk.

$$\nabla^2 \psi = -\frac{\rho}{\epsilon_s} \quad (4.10)$$

$$\rho(x, y) = q\left[p(x, y) - n(x, y) - N_A\right] \quad (4.11)$$
The terms in 4.11 for electron and hole concentration are well defined and given by 4.12 and 4.13 respectively. The below equations are given for a four terminal device and therefore the potential applied to the Source and Drain has be be accounted for and this is given by \( V(x) \). Furthermore, we can see that both electron and hole concentrations a function of the Surface Potential and therefore Gate potential, in addition the Source to Drain potential also plays a role. The fermi level is also accounted for in electron and hole concentration and is given by \( \phi_{ms} \).

\[
n(x, y) = N_A e^{[\psi(x,y) -(2\phi_F - V(x))] / \phi_t} \tag{4.12}
\]

\[
p(x, y) = n_i e^{[\phi_F - \psi(x,y)] / \phi_t} \tag{4.13}
\]

Using 4.12, 4.13, and a few other simple well known semiconductor relations, we can derived an expression for the channel charge (48). This is given by 4.14 this equation also assumes a few approximations which simplify the equation. One of the principal ones being the gradual channel approximation (GCA) (50). The GCA approximation considers most of the change in surface potential to occur between the bulk and the tunnel layer interface which is along the y axis, while there is little change along the channel, or the x axis. Therefore for we can ignore the partial derivative of the surface potential along the x axis, which reduces 4.11 to 4.14.

\[
\frac{\partial^2 \psi_s}{\partial y^2} = -\frac{q N_a}{\epsilon_s} \left[ e^{-\psi_s(y)/\phi_t} - 1 - e^{-2\phi_f'/\phi_t} (e^{-\psi_s(y)/\phi_t} - 1) \right] \tag{4.14}
\]
Equation 4.14 can be simplified further using integration and boundary conditions such as that the surface potential will go to zero deep in the bulk, $\psi_s = 0$ as $y \to \infty$. With these methods and conditions applied, we can derive a term for $Q_c$, which is given by 4.15.

$$Q_c = \pm \sqrt{2q\varepsilon_s N_A} \left[ \phi_t e^{-\psi_s/\phi_t} + \psi_s - \phi_t + e^{-2\phi_f/\phi_t} \left( \phi_t e^{(\psi_s-V(x))/\phi_t} - \psi_s - \phi_t e^{-V(x)/\phi_t} \right) \right]^{1/2} \quad (4.15)$$

At this point, most of the variables in 4.9 are known, with $Q_o$ a constant for trapped charges in the oxide. With some algebra, we can combine 4.9 and 4.14 to get the transcendental equation of MOSFET given by 4.16.

$$V_{gb} - V_{fb} - \psi_s = \gamma \sqrt{\psi_s + \phi_t \left( e^{-\psi_s/\phi_t} - 1 \right) + \phi_t e^{-2\phi_f-V(x)/\phi_t} \left( e^{\psi_s/\phi_t} - 1 \right)} \quad (4.16)$$

The power of the transcendental equation of MOSFET is that it is an all region physics based model which requires no fitting parameters or diving into regions of operation for modeling. To make 4.16 simpler for computation, often it is divided into regions of operations where it reduces into simpler form but loses accuracy, often fitting parameters are introduced to compensate.

There is no closed form solution to 4.16 as it is a coupled equation and cannot be isolated for $\phi_s$, it must be solved numerically. Fortunately there are efficient algorithms that can solve this equation quickly (51). Often 4.16 is written as 4.17 for computation purposes.
\[ f = \frac{1}{\gamma^2} (V_{gb} - V_{fb} - \psi_s)^2 + \psi_s + \phi_t \left( e^{-\psi_s/\phi_t} - 1 \right) + \phi_t e^{(-2\phi_f - V(x))/\phi_t} \left( e^{\psi_s/\phi_t} - 1 \right) = 0 \] (4.17)

The numerical method employed for solving 4.16 is to use Newton-Raphson methodology shown in 4.17. The basic methodology of 4.18 is to guess a value for \( \phi_s \) and then evaluate 4.17 divided it by its first and second derivative and subtract it from the guessed value for \( \psi_s \). The new value for \( \psi_s \), found on the right hand side of 4.18, is then reinserted into the 4.18 as the new guess and 4.18 and its derivatives are reevaluated. With each iteration, the subtracted term will approach zero so a stable \( \psi_s \) will be found. This iteration is usually done from three to five times for a good convergence for \( \phi_s \). It should be noted that (17) is a slightly modified version of Newton-Raphson as denominator term has more components than just first derivative of 4.17, this is done for faster convergence.

\[ \psi_s = \psi_s - \frac{f}{f''} \quad (4.18) \]

Getting a good convergence for \( \psi_s \) requires several iterations and a good initial guess, this guess is usually provided by a basic region based surface potential equation which is then inserted into 4.18 for iteration. A good initial value is prerequisite as 4.18 will converge on a bad value, therefore a basic understanding of surface potential profile as a function of Gate and Drain voltage must be understood. Figure 4.4a and b, shows the initial guess, in blue, and the convergence of surface potential value. Figure 4.4, shows how the initial guess difference from the converged or actual value. The iterative step requires much more computation time than other approaches, which is why some models avoid the iterative step and simply follow the region based approach with fitting parameters. The need for fitting parameters can also be observed in Figure 4.4.

Once surface potential has been calculated, the drain current can be calculated by using
the Poa-Sah current equation (47), given by 4.19 and 4.20. The Poa-Sah equation uses two integrals, one integration is done for surface potential going from the bulk region to the surface and the other for the voltage from Source to Drain. Unfortunately, this also proves to be computationally intensive as it involves several nested loops to evaluate the inner terms of the integral before numerical integration can be done.

\[
I_{DS} = \frac{W}{L} q N_A \int_{V_{SB}}^{V_{DB}} \mu \int_{0}^{\psi_s} e^{(\psi - 2\phi_f - V(x))/\phi_t} \frac{E(\psi)}{E(\psi)} d\psi dV
\]

(4.19)

\[
E(\psi) = \frac{Q_c}{\epsilon_s}
\]

(4.20)

The surface potential model can now be modified to include the novel behaviors of QDSL based transistors and memory. The MOSFET transcendental equation, 4.16, can be modified by accounting for the charge storage capabilities of the QDSL to produce intermediate states.
4.2 New Formalism for Surface Potential Model Incorporating QDSL

To incorporate QDSL charge into the surface potential model, one needs to incorporate the QDSL layer charge \( Q_{FG} \) into the model. As previously developed and restated here, the QDSL charge storage capability is defined as 4.21. This incorporates the density of states of the mini-bands and their filling as a function of applied voltage which is represented by the fermi level.

\[
Q_{FG} = \int_{w_1} f(E)g_{w_1}(E) + \int_{w_2} f(E)g_{w_2}(E) + \int_{w_3} f(E)g_{w_3}(E) \quad (4.21)
\]

This \( Q_{FG} \) is then incorporated back into charge conservation from equation 4.4 and modified to produce 4.22. In addition the \( \phi_{MS} \) must also incorporate the additional charge found in the QDSL, equation 4.6 is rearranged to put it in terms of flatband voltage \( V_{FB} \) as shown in 4.23. This is done because \( V_{FB} \) is no longer a constant, unlike in most circumstances for a MOSFET, it can be calculated once and held as a constant. By incorporating the \( Q_{FG} \), the \( V_{FB} \) become a function of the Gate voltage \( V_{gb} \). The \( V_{FB} \) is no longer a constant but rather a dependent variable to the independent \( V_{gb} \) due to the \( Q_{FG} \) being a function of Gate voltage which is expressed as fermi level energy.

\[
Q_g + Q_o + Q_c + Q_{FG} = 0 \quad (4.22)
\]

\[
V_{FB} = \phi_{MS} + \frac{Q_o}{C_{ox}} + \frac{Q_{FG}}{C_{ox}} \quad (4.23)
\]
It now becomes straight forward to apply it to the the transcendental equation of MOSFET, 4.16, must now incorporate a flatband voltage dependent on the Gate voltage, as seen in equation 4.24. This this modification to $V_{FB}$ incorporates a whole host of new challenges as the 4.16 had only one dependent variable, that of surface potential $\psi_s$, but now there is two with $V_{FB}$ being the other.

$$V_{gb} - V_{fb}(V_{gb}) - \psi_s = \gamma\sqrt{\psi_s + \phi_t \left(e^{-\psi_s/\phi_t} - 1\right) + \phi_t e^{(-2\phi_t - V(x))/\phi_t} \left(e^{(\psi_s/\phi_t)} - 1\right)}$$ (4.24)

The new transcendental equation for QDSL based devices can now be solved using the Newton-Raphson method and modified procedure outlined in (51). Several major changes must be incorporated to make the computation efficient, one is to calculate QDSL charge as a function of Gate voltage prior to solving the 4.24, this has been previously done and shown again here in Fig. 4.5. The rational for computing the QDSL charge once and keeping it as an accessible data set is to save time as this calculation is time consuming and importantly not necessary for on the fly calculation. The initial QDSL charge, $Q_{FG}$, can be calculated for a nominal unit area and any device variation will be reflected at the Drain current with W/L being the critical adjustment factor.
4.2.1 Modified Surface Potential Results for MVL Devices

Using this new approach, with several major assumptions made that are discussed further down, a QDSL based MVL transistor and NVM can be simulated. Figure 4.6 shows Silicon QDSL based devices for which results are computed for. These results are preliminary and author must caution that their accuracy remain unverified and presumably low but the trend to capturing these unique device behaviors shows potential. A direct comparison should not be made between the simulated results and fabricated device results as simulated results did not take into account all the different parameters necessary for a true comparison. More importantly the model need much more refinement prior to any real comparison that can be made.

Figure 4.7a shows matlab calculation using equation 4.24 equation and data set from 4.5 for a Silicon devices shown in Fig. 4.6a with an intermediate step clearly visible in the Drain current. This is then compared to 4.7b, which has current values from a fabricated device.
The simulated results for MVL transistors had a few key assumptions made, one of them is that a certain voltage charges would start tunneling into the QDSL layer. This assumption was made using simple hand calculation but would need to be formalized for a more accurate results. In particular a quantum tunneling model could be made which would calculate the tunneling current as a function of oxide thickness and Gate voltage.

Secondly, the assumptions made in (51) for $\psi_s$ were modified to saturate around the where the intermediate state is expected, this was done so that the new transcendental equation
can converge to the right values for surface potential. This is necessary as the transcendental equation is a non-unique and can produce inaccurate results if initial guess for Newton-Raphson is not close. A lot of simple calculation were made as to where the intermediate state would appear, a much more rigorous model must be devised and implemented for better model accuracy.

4.2.2 Modified Surface Potential Results or QDGNVM Devices

Equation 4.24 can also simulate results for QDG memory devices, as in Fig. 4.6b. It is not necessary to alter any equation as for MVL to NVM but to make the assumption that there will be no leakage to the Gate and that an independent Write pulse will be applied to charge the QDSL layer. The way a bit of memory is stored is determined by a threshold shift, $V_{th}$, this shift occurs due to charges present in QDSL layer. Equation 4.25 is the standard definition of threshold voltage when charge conservation is incorporated, with the $\psi_s$ and $\psi_{FG}$ calculated by dividing $Q_c$ and $Q_{FG}$ by oxide capacitance.

$$V_{th} = V_{FB} + \phi_o + \gamma \sqrt{\phi_o} \quad (4.25)$$

A bit of memory can be shown to be stored by first calculating a conventional transistor $I_D - V_g$, which is known as a Read cycle. This should give a conventional transistor value, with a pulse applied, known as a Write pulse, a certain amount of charges will tunnel into the QDSL layer. The amount of charges in the QDSL layer will be a function of the Write pulse voltage and duration, the first can be calculated from Fig. 4.5 with modifications taken into account for formation of the inversion layer in the channel. Once a Write pulse is applied and charges are present in QDSL layer then a Read cycle is again initiated where a lateral shift in $I_D - V_g$ is seen, this is shown in Fig. 4.8. The charges present in $Q_{FG}$ alter the $V_{th}$
by changing the $V_{FB}$ as shown in 4.25.

Figure 4.8: Silicon QDG memory showing a) simulated results with threshold voltage shift and b) actual fabricated device results

The simulated results of the NVM charge storage can be seen in Fig. 4.8a where the arrow indicates the difference between the pre and post Write cycle, which shows a threshold shift. Figure 4.8b (52) fabricated device results of an NVM, this is done not for direct comparison purposes but to demonstrate the general trend alignment between the simulated result and real devices. The simulated NVM results correlates well with fabricated results in that a noticeable $V_{th}$ is produced. It should be noted that simulated result does not show the full hysteresis loop as seen in the fabricated devices due to numerical calculation often break when converging to zero, this could be fixed by further refinement of the code.

A note of caution must added here as the simulated results are a first iteration of what must be further code refinement which captures actual device physics. The intent of the first order simulation is to work out the device physics and show its plausibility. Several key steps were not taken, one of which is the voltage at which current beings to tunnel into the layer and the rate at which it does dictates charging time of the QDSL. The tunneling voltage is calculated by hand and the charging time is skipped, the charging time would ostensibly dictate how short in duration the Write Cycle can be and it should be set several
times this value. The above information is critical but is beyond the scope of this study as it would require quantum simulation. A second important step is skipped in that a code can be written with the QDSL charge being a function of the Write pulse, this again is substituted with hand calculation and its corresponding variances. The intent is not to develop a high fidelity model but rather demonstrate the validity of the new QDSL based Surface Potential Formalism.

4.3 Discussion

Modeling remains a key component for today’s understanding and implementation of transistor devices. To understand QDSL based devices and their MVL and NVM behavior a modified Surface Potential numerical model is developed and then first order simulation conducted. The simulated results show the correct trends when compared to actual fabricated devices but require further refinement for a proper comparison. Certain issues such as speed of numerical calculation, high fidelity, calculating resistance and capacitances, and incorporating quantum effects into the model must be addressed. Some of these can be addressed quickly, such as resistance and capacitance since surface potential profile is known while others like quantum tunneling of the insulating layer will take significant effort.

One of the primary intent of undertaking a physics based mathematical model of QDSL based devices is to allow circuit designers ability to design and simulate unique MVL or NVM circuits. These circuits can then show higher processing ability while keeping current CMOS designs and processes. Circuit designs often take the first steps in a technology change as they must be able to design and demonstrate complex circuit functionality prior to technology adaption. It is hoped that with a MVL and NVM numerical model can further QDSL based device implementation.
Chapter 5

Conclusion

Current CMOS technology may not be sustainable as transistor technology is continuously scaled (53). Thus far the collective semiconductor industry has been able to scale down to below 100 nm using a high-k dielectric as Gate insulators instead of silicon oxide which help in mitigating leakage current to the gate by using a higher permittivity material which allows for a larger thickness compared to silicon oxide (54). The industry has been further able to scale down below 20 nm by switching from planer transistor to trigate configuration called FinFET, where the width of the transistor is limited and metal Gate surrounds the channel on all three sides (55). These two fundamental shifts have allowed transistors to keep shrinking while a many other improvements have been taken to overcome short channel effects such as punch through, Drain Induced Barrier Lowering, and sub-threshold leakage current to name a few (9). There are even transistors being designed for production at 5nm channel lengths (56), this represents 10 time the lattice constant of Silicon at 0.543 nm. Although previous claims of the end of transistor scaling have been proven very wrong, at what point does manufacturing cost and quantum limits become so great that it is no longer productive to go smaller any longer.

As stated before, there exists many candidate which can replace the current CMOS technology but most often they involve totally new processes of fabrication and or exotic
materials. Although some of these solution may prove viable, QDSL based devices prove promising due to their CMOS pedigree, being consistent with all current CMOS processing and the ability to deliver higher processing capability by using high logic and has demonstrated higher memory density by storing 2 bit in one transistor all while keeping transistors relatively large which avoid short channel effect. The QDSL based devices should also be scaled down but the increased processing power may offset this requirement.

We have demonstrated viability of MVL logic based devices which were simulated using SWS-FET, these SWS-FET can also be fabricated out of QDSL structures and efforts are currently underway. The MVL based Multiplexer and De-multiplexer have shown significant transistor reduction while maintaining same functionality. The reduction in number to transistor can lead to significant area saving and reduction in cost. Most likely the area savings will be used to put in more MVL devices thus further increasing processing power.

The fabricated QDSL based MVL and multi-bit NVM transistors, while showing novel characteristics, required refinement for applications in real world environment as their electrical characteristics did not have significant noise margins required. A solution was shown for this issue by increasing the QDSL layers from two to four with a mixed configuration where both Germanium and Silicon QDSL are used which is predicted to provide significant noise margin. Novel four layer fabrication methodology is devised and proven to allow refinement of MVL and NVM devices. This steps are necessary as real work integrated circuits are built to certain margins and must have adequate tolerances. With four layer the QDSL application become more realizable.

Finally a mathematical model has been devised for QDSL based transistor for both MVL and NVM using a modified high fidelity Surface Potential model. The key physics implication of using QDSL is that the $V_{FB}$ become a function of $V_g$, as the QDSL retains charge with increasing Gate voltage it shifts the flat-band voltage along with it producing intermediate states in the case of MVL and threshold voltage shift in the case of memory. The mathematical model is then simulated and trends compared to actual fabricated devices
to show an overall good correlation. The modified Surface Potential model can give designers and insight into working of the QDSL device and allow them to produce actual circuits which incorporate their unique quantum behavior.

It is hoped that by demonstrating the viability of QDSL based devices and their advantages in fabrication, enhanced computational power, and higher logic and memory behavior that they can be one of the answers to the semiconductors industries perennial challenge with scaling.
Chapter 6

References

Bibliography


[51] Rafael Rios, Sivakumar Mudanai, Wei-Kai Shih, and Paul Packan. An efficient surface...


