Modeling and Design of Nanoscale Ferroelectric and Negative-Capacitance Gate Transistors

by

Yu-Hung Liao

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Sayeef Salahuddin, Chair
Professor Chenming Hu
Professor Junqiao Wu

Fall 2021

Modeling and Design of Nanoscale Ferroelectric and Negative-Capacitance Gate Transistors

Copyright 2021
by
Yu-Hung Liao

Abstract

Modeling and Design of Nanoscale Ferroelectric and Negative-Capacitance Gate Transistors

by

Yu-Hung Liao

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Sayeef Salahuddin, Chair

Integration of Hf-based ferroelectric materials into the MOSFET gate stack introduces new physics that potentially enables continued scaling for both CMOS and memory devices. The electrostatic potential amplification effect arises when the device is engineered to depolarize and hence stabilize the ferroelectric negative capacitance state, which eliminates the ferroelectric hysteresis and enhances the gate control. On the other hand, when the ferroelectric layer is thick enough to overcome the depolarization field, non-volatile memory function can be realized through the remnant polarization states. Such systems exhibit process and structural compatibility with CMOS technology and thus are of significant application relevance. Experiments have demonstrated low-voltage operations for both types of devices, showing promising prospects for next-generation logic devices with high reliability.

To harness the effects for designing the NCFET and FeFET devices, it is important to understand the behaviours of the ferroelectric layer in the relevant states. In this dissertation, the polarization responses of negative capacitance states are modeled from different aspects through self-consistent device and circuit models that are calibrated to experimental measurements. First, the 101-stage ring oscillator consisting of 14 nm FinFET devices with a ferroelectric gate layer that exhibits negative capacitance are used to study the response speed of the negative capacitance effect. The consistency of the device DC characteristics and the circuit oscillations with less than 10ps per-stage delay confirm the fast negative capacitance response. Next, NCFET device characteristics and scaling trends that are unexpected by conventional theory were explained with non-linear negative capacitance responses as described by the Landau's phenomenology. Design insights are drawn, and the observation bodes well for extending the MOSFET gate length scaling limit. Finally, high-speed NCFET RF operations are projected and designed based on experimental characterization results and Monte-Carlo transport simulations. Silicon-channel NCFETs are expected to achieve cutoff frequency over 400GHz with standard SOI device structures and over 650GHz with air-gap spacers.

The carrier dynamics of memory program and readout operations for an n-type SOI FeFET are studied through TCAD simulations. Gate-induced drain leakages during the program operation with large negative gate biases can result in excessive hole concentrations with lifetime exceeding microseconds during the hold phase. While the excess holes result in approximately 100ps delay in the readout phase before a large read margin can be detected, it benefits the program speed and the quasi-steady-state read margin. Therefore, such effects can be utilized for low-power memory applications through proper SOI FeFET designs with the awareness of excess carrier dynamics.

To my family.

Contents

Contents ii

List of Figures iv

List of Tables xi

1 Introduction 1
1.1 Needs For Electrical Scaling of CMOS Gate Stack 1
1.2 Ferroelectric Negative Capacitance 3
1.3 Negative Capacitance Transistors 11
1.4 FeFET Memory 13

2 Response Speed of Negative Capacitance FinFETs 14
2.1 Introduction 14
2.2 Validation Methodology 15
2.3 Experiments and Models 18
2.4 Conclusions 23

3 Modeling of the Anomalously Beneficial Gate-Length Scaling Trend 24
3.1 Introduction 24
3.2 Sub-threshold Scaling Trends 24
3.3 Modeling of the NCFET Scaling Trends 26
3.4 Simulation Results 29
3.5 Conclusion 33

4 Modeling of the Near-Threshold Swing Reduction Effect 34
4.1 Introduction 34
4.2 Device Characterizations 35
4.3 TCAD Model 39
4.4 Discussion 42
4.5 Conclusion 46

5 Performance Modeling of Quasi-Ballistic NCFET 47

5.1 Introduction . . . 47
5.2 Device Structure . . . 48
5.3 Modeling Methodology . . . 50
5.4 Results . . . 55
5.5 Conclusion . . . 68

6 Dynamical Effect of Excess Carriers On SOI FeFET Memory Operations 69
6.1 Introduction . . . 69
6.2 FeFET Design and Simulation Model . . . 70
6.3 Results and Observations . . . 71
6.4 Design Perspectives . . . 78
6.5 Conclusion . . . 81

7 Conclusions 82
7.1 Chapters Summary . . . 82
7.2 Future Works . . . 84

Bibliography 86

List of Figures

1.1 Schematic (a) $\mathcal{G}_{FE} - P_{FE}$ relationships at different E_{FE}. Possible steady states are marked. For the NC states, P_{FE} decreases as E_{FE} increases in contrast to classical trend. However, they are not stable in a bulk system because they correspond to local maxima of $\mathcal{G}_{FE}$. (b) $E_{FE} - P_{FE}$ (solid) and $D - P_{FE}$ (dashed) relationships. Both have negative slopes for small $|P_{FE}|$, corresponding to the NC state. (c) Illustration of different fields for an NC state. P_{FE} responds in the same direction as E_{ext}. However, the depolarization field generated by P_{FE} ($E_{dep} = -P_{FE}/\epsilon_b$) is stronger than E_{ext}, reversing the direction of the net electric field E_{FE}. 4
1.2 (a) Illustration of the fields in the FE-DE stack with applied voltage when FE is in the NC state. The displacement D is continuous throughout the layers, which connects E_S and E_{FE} to V and P_{FE}. (b)(c) The energy landscapes of a FE-DE stack system under different applied voltages. The system satisfies condition 1.20, so G is stabilized at the low-P_{FE} state, and the hysteresis is eliminated. The overall capacitance of the system is positive, so as V increases, Q and D increases, which drives P_{FE} in the same direction, while E_{FE} moves in the opposite direction due to the NC effect. 5
1.3 Relationship between the cation displacement and the forces on the cation atom. The net force is decomposed into internal terms that are unrelated to the applied field and the E_{FE} term. (a) is without a DE layer and V=0, so $E_{FE} = 0$ regardless of z_B and P_{FE}. The internal terms make the NC state unstable, while only the remnant polarization states are stable. (b) and (c) are with a DE layer that is thick enough to stabilize the NC state which becomes the only steady state. The FE polarization induces a depolarization field that restores the cation from fluctuations. When the voltage bias is changed from (b) to (c), the dashed curve shifts vertically, which results in a new steady state. 8
1.4 (a) Simulated Q-V results for a DE capacitor, an FE capacitor, and a stack structure combining both. LK model is used to simulate the P_{FE} responses. (b) A snapshot of the electric field profile for the FE-DE stack structure in NC state. 10
1.5 NCFET modeling methodology flow with controlled experiments. 13

2.1 Schematics of the 101-stage ring oscillator when it is disabled (left) for IDDQ measurements and when it is enabled (right) for IDDA and delay measurements. Fan-out loads are not shown. 16

2.2 (a) V_{th}-matched I_D-V_G characteristic of Control MOSFET and NCFET. For comparison, simulated characteristics of devices with 3 nm DE is included. (b) A schematic 'S' curve of the ferroelectric showing the region of operation. 18
2.3 Measured and modeled I_D-V_G for (left) an n-type NCFET and for (right) a p-type NCFET. Both semi-log scale and linear scale current are plotted. Dotted and solid lines represent measurements and modeling, respectively. 19
2.4 Measured and modeled (left) I_D-V_D at various V_{GS} and (right) C-V at $V_{DS} = 0V$ for both the NC-NFET and NC-PFET. 19
2.5 Simulated and measured (a) IDDQ and (b) IDDA on the 101 stage FO3 RO. The simulation was performed with quasistatic models calibrated to the DC measurements. 20
2.6 (a) Operation range corresponding to a V_{DD} where the PFET remains in NC regime. (b) Simulated and measured IDDA at the V_{DD} indicated in (a). 21
2.7 The simulated and measured (left) C_{eff} and (right) gate stage delay τ for the FO3 NCFET RO. (Left) The dashed line is simulated without wire load, and the solid line is simulated with 0.4fF wire load. (Right) The delay is reproduced by the model with 0.4fF wire load. 22
2.8 (Left) Simulated and measured gate stage delay. The shaded area stands for the V_{DD} range under which the device model is calibrated to DC measurements. (Right) Simulated gate stage delay and the projection with matched V_{th}. The shaded area stands for the trustworthy regime for matched-V_{th} simulations. . . 22
2.9 (a) Measured C_{eff} and R_{eff} from 4-fin NCFET ring oscillator. Behavior of C_{eff} and R_{eff} follow expected trends. (b) Delay vs. IDDA curves for 2 fin and 4 fin NC-FinFET ring oscillators. Stage delay as small as 7.2 ps was measured. . . . 23

3.1 (a) Electrostatic potential profile and (b) current density profile of a short-channel n-type SOI MOSFET biased in sub-threshold regime ($V_{GS} = 0V$ and $V_{DS} = 0.05V$.) Because of the built-in potential between the gate and the source and drain, the electrostatic coupling causes the potential to be lower at the channel surface than at the bottom. As a result, the leakage is dominated by the bottom path, which implies the TOB best describes SS and DIBL is located at the bottom of the Si layer. 25
3.2 Extracted (a) SS scatter plot, (b) control-to-NC $m - 1$ ratio with error bar, (c) DIBL scatter plot, and (d) control-to-NC DIBL ratio with error bar, measured for different L_g devices. Each data point in (a) and (c) corresponds to one measured device, where a small horizontal shift is used to separate control and NC results for better visualization. The bars in (b) and (d) indicate the sample standard deviation of each ratio due to device-to-device variations. 27

3.3 Electric field profile of a short channel SOI MOSFET in sub-threshold regime. The light color in the silicon channel near the source and the drain edge is due to the inner fringing field coupling, and it does not fully decay at the middle of the channel. (b) is a zoomed-in view of (a) from mid-channel region to the drain edge with arrows indicating the field direction. The electric field on the gate stack is nearly vertical for the mid-channel region. 28

3.4 Simulated (a) SS, (b) $m-1$ ratios of control devices to improved devices, (c) DIBL, and (d) DIBL ratios of control devices to improved devices, as a function of L_g. The control devices consist of 2.31nm-EOT gate stacks. Improved devices include "higher-κ" conventional devices (EOT=1.73nm) and NC devices ($\alpha = -2.4 \times 10^{11} cm/F$, $\beta = 0 cm^5/(FC^2)$, $\gamma = 3.2 \times 10^{33} cm^9/(FC^4)$, $\epsilon_r = 19$, and SiO_2/FE interface charge density of $-0.64 \mu C/cm^2$). 30

3.5 Simulated out-of-plane electric field (E_y) in the interfacial layer of NCFETs with different L_g as a function of position along the channel (x) from source edge ($x = 0$) to drain edge ($x = L_g$) at $V_{GS} = 200mV$ and $V_{DS} = 50mV$ (in sub-threshold regime.) 31

3.6 Simulated (a), (c) $Q-V$ relations of the ferroelectric layer and (b), (d) $C-Q$ relations of the gate stack at the mid-channel of NCFETs with different L_g. For each L_g, the gate voltage is ramped at which I_D is from 10pA/μm to 100pA/μm at V_{DS}=50mV. For (a) and (b), the ferroelectric parameters are $\alpha = -2.4 \times 10^{11} cm/F$, $\beta = 0 cm^5/(FC^2)$, $\gamma = 3.2 \times 10^{33} cm^9/(FC^4)$, and $\epsilon_r = 19$. The devices in (a) and (b) are simulated with $-0.64 \mu C/cm^2$ fixed charges at the SiO_2/FE interface. For (c) and (d), the ferroelectric parameters are $\alpha = -1.4 \times 10^{11} cm/F$, $\beta = -1.2 \times 10^{22} cm^5/(FC^2)$, $\gamma = 8.0 \times 10^{32} cm^9/(FC^4)$, and $\epsilon_r = 19$. No interface charges are assumed in (c) and (d). 32

3.7 Simulated (a) $m-1$ ratios (b) DIBL ratios between control MOSFETs and ferroelectric FET with nonlinear positive-capacitance oxide. For the imaginary positive-capacitance ferroelectric material, $\alpha = 4 \times 10^{11} cm/F$, $\beta = 10^{22} cm^5/(FC^2)$, and $\gamma = 10^{31} cm^9/(FC^4)$. 33

4.1 (a) Schematic cross-section of the SOI n-MOSFETs. Both HfO_2 (Control) and HZO (NC) devices have the same geometry. (b) SS of HfO_2 and HZO devices demonstrating conventional and unconventional trend with respect to I_D, respectively. (c) Equivalent circuit diagram of the MOSFET in sub-threshold regime. 35

4.2 Demonstration of inversion symmetry breaking and switchable electric polarization in 3 nm ZHO film. (a) Synchrotron in-plane grazing-incidence (IP-GiD) diffraction indexed to the polar orthorhombic phase ($Pca2_1$) associated with ferroelectric order in HfO_2-based systems. (b) Piezoresponse force microscopy (PFM) phase-contrast images demonstrating stable, bipolar, remanent polarization states that can be overwritten into opposite polarization states. 36

4.3 Measured and TCAD-simulated SS-I_D relations for Control and NC n-MOSFETs of four gate lengths at V_{DS} =50mV. Applied V_{gs} for measurement is -0.4V ∼ 1V. Multiple devices are measured for each gate length, all of which exhibit negligible hysteresis. 38

4.4 Measured and TCAD-simulated SS-I_D relations for Control and NC n-MOSFETs of four gate lengths at V_{DS} =500mV. Applied V_{gs} for measurement is -0.4V ∼ 1V. Multiple devices are measured for each gate length, all of which exhibit negligible hysteresis. 39

4.5 (a) SS averaged over I_D=0.1∼1nA/μm at V_{DS}=0.05V for measured and TCAD-simulated Control and NC MOSFET. Each marker presents one device having the corresponding gate length. (b) Experimentally estimated SS improvement for different gate lengths. Each error bar presents the estimated mean and one standard deviation. (Inset) TCAD-modeled Improvements for the four gate lengths. . . . 40

4.6 (a) Measured and TCAD-simulated C-V for p-type doped MOS capacitors with HfO_2 and ZHO gate stacks. No frequency dispersion is observed up to 300kHz for measurement gate bias range. (b) Constant-current (I_D=10nA/μm) threshold voltage (V_t) at V_{DS}=0.05V for measured and TCAD-simulated Control and NC MOSFET. Each marker presents an extraction for one device. Experimental V_t are 0.18V higher than reference simulations with nearly neutral net oxide and interface trapped charge. The discrepancy is resolved when -0.67μC/cm^2 fixed charge is added to the model. 40

4.7 (a)-(b) Measured and TCAD-simulated I_D-V_G relations for Control and NC n-MOSFETs with L$_G$=30nm at V_{DS}=0.5V. Each measurement curve characterizes one MOSFET. (c)-(d) Simulated SS-ID from transient gate-sweep TCAD simulations (L$_G$=50nm.) The forward and reverse gate voltage sweeps produce overlapping curves, implying zero hysteresis. 42

4.8 Extracted relations between external charge and net electric field for the polar layer characterized at the mid-channel of NCFETs at V_{DS}=50mV. For each gate length, the gate voltage is ramped at which I_D is from 0.1nA/μm to 1nA/μm. . 43

4.9 Simulated relations between drain current and (left axis) local gate charge density / (right axis) local external electric field at (a) V_{DS}=0.05V and (b) V_{DS}=0.50V for NCFETs. For each gate length, the solid line is extracted at mid-channel gate stack region, and the dash line is extracted at the region 15nm laterally from mid-channel toward the drain. The white-background region corresponds to bias conditions at which the polar layer exhibits negative capacitance, while the dark background stands for positive-capacitance regime. 43

4.10 Relations between small-signal capacitance-equivalent SiO_2 thickness of the gate stack and applied external electric field extracted at I_D=0.1∼1nA/μm. Extractions are made at different lateral positions under two drain biases. 44

4.11 Simulated (a) SS and (b) DIBL for various L$_G$ with the Control gate stack and a charge-engineered gate stack that exhibits sustained NC behavior. FinFET structures with 7nm fin thickness are assumed according to the IRDS “5nm” node. 45

5.1 Overview of the front-end process flow in this study. The shown example corresponds to the "faceted RSD" flow as described in [85], but is gate-last to accommodate the low thermal budget of the ferroelectric-metal gate stack. 48

5.2 Three types of devices studied in this work. (a) The typical RSD structure with nitride spacer and nearly-vertical RSD side walls. (b) The air-gap spacer structure where the nitride spacer is removed from (a). (c) The slanted-side-wall RSD structre formed through specialized epi-Si:P module. 50

5.3 (a) The equivalent circuit diagram for modeling small-signal responses of the SOI FET device. (b) Decomposition of the SOI FET into the intrinsic channel and the RSD subsystems. (c) Simulated structure for extracting intrinsic components g_m, r_o, C_{gsi}, and C_{gdi}. (d) Simulated structure for parasitic components associated with the raised source (C_{gse1}, C_{gse2}, and r_s) and the drain (C_{gde1}, C_{gde2}, and r_d.) 51

5.4 Simulated (a) I_D-V_{GS} and (b) g_m-V_{GS} at V_{DS}=0.7V of the same device structure through different transport solvers with similar low-field mobility and optical phonon scattering models. The drift-diffusion simulation produces lower current than the more rigorous solvers. The single-particle MC simulation produces a slightly higher current than the MC simulation with self-consistent electrostatic potential. 53

5.5 (a) The measured and MC-simulated low-field electron mobility for long channel NCFET devices. (b) The simulated long-channel capacitance density. 53

5.6 MC steady-state profile of (a) conduction electron density and (b) electron velocity in the transport direction for a 4nm-spacer NCFET at $V_{GSi} = 0.45V$ and $V_{DSi} = 0.7V$. (c) The sheet density of the right-going electrons (N_{e+}) and the left-going electrons (N_{e-}.) (d) The averaged velocity in the transport direction for the right-going electrons (v_{ey+}) and the left-going electrons (v_{ey-}.) The x-axes for (c) and (d) represents the lateral position from the inner source (-14nm) to the inner drain (14nm). 56

5.7 (a) The sheet conduction electron density (N_e,) (b) the local capacitance density characterizing the response of the N_e to the change of V_{GSi}, and (c) the averaged electron velocity in the transport direction (v_{ey}) at $V_{GSi} = 0.4V$ and $0.5V$. Results from different transport simulators (DD and MC) are shown. (d) The v_{ey} and the ratio of the N_{e+}/N_e at $V_{GSi} = 0.45V$. The device is the same as that shown in Fig. 5.6, and the position follows the same definition. 58

5.8 Simulated (a) I_D-V_{GSi}, (b) g_{mi}-V_{GSi}, (c) $(C_{gsi} + C_{gdi})$-V_{GSi}, and (d) f_{ti}-V_{GSi} at V_{DSi}=0.7V for faceted RSD devices with various spacer thicknesses and dopant diffusion times. All subplots share the same legends. Small fluctuations of g_{mi} and f_{ti} are due to the random nature of MC simulations. 59

5.9 Simulated (a) I_D-V_{GSi}, (b) V_{DSi}/I_D-$1/V_{ovi}$, and (c) the extracted series resistance for the intrinsic device at V_{DSi}=0.05V for faceted RSD devices with various spacer thicknesses and dopant diffusion times. The legend for (a) also applies to (b). . 60

5.10 The vertically-averaged conduction band energy for the SOI versus the lateral position from inner S to inner D at various V_{GSi} for (a) V_{DSi}=50mV and (b) V_{DSi}=0.7V. Vertical dashed lines indicate the boundary for the gate stack and the spacers. The simulated devices have 8nm spacers undergoing 2min (dashed curves) or 10min (solid curves) dopant diffusion at 800°C. The 2min-diffusion cases have the top of the potential barrier in the spacer extension region at the source side and hence much weaker gate control of the barrier height as compared to the 10min-diffusion cases. 61
5.11 Extract (a)-(c) I_D at V_{GSi}=0.45V and V_{DSi}=0.7V, (d)-(f) g_{mi} at at V_{GSi}=0.45V and V_{DSi}=0.7V, and (g)-(i) SS at V_{GSi}=0V and V_{DSi}=0.7V. The same column of subplots represents the same type of device structure as indicated in (a)-(c). . . 62
5.12 Extract (a)-(c) C_{gsi} and C_{gdi} at V_{GSi}=0.45V and V_{DSi}=0.7V, (d)-(f) f_{ti} at at V_{GSi}=0.45V and V_{DSi}=0.7V. The same column of subplots represents the same type of device structure as indicated in (a)-(c). 63
5.13 Extract (a)-(c) C_{gse1}, (d)-(f) C_{gse2}, (g)-(i) r_s, and (j)-(l) the voltage drop on the RSD for the source side at V_{GSi}=0.45V and V_{DSi}=0.7V. The same column of subplots represents the same type of device structure as indicated in (a)-(c). . . 65
5.14 Extract (a)-(c) C_{gde1}, (d)-(f) C_{gde2}, and (g)-(i) r_d, and (j)-(l) the voltage drop on the RSD for the drain side at V_{GSi}=0.45V and V_{DSi}=0.7V. The same column of subplots represents the same type of device structure as indicated in (a)-(c). . . 66
5.15 Extracted f_t at at V_{GSi}=0.45V and V_{DSi}=0.7V for various device structures and process parameters. Legends are shared by all the subplots. The same column of subplots represents the same type of device structure as indicated in (a)-(c). . . 67

6.1 2D schematic of the TCAD-simulated SOI-FeFET device. 70
6.2 I_D-V_G readout after different zero-bias-hold times. Solid: high-V_T phase. Dashed: low-V_T phase. The rapid waveform ramp time of 5 ns V^{-1} ensures no polarization switching occurs during the readout. The magenta curve shows slightly different polarization values than the other curves due to a short hold time which is insufficient for the ferroelectric to reach its equilibrium polarization at $V_{GS} = 0$ V. 72
6.3 (a) Band energies (blue) and hole band-to-band generation rate (green) of a horizontal cross-section taken in the silicon channel (Y= 11 nm, 1 nm from the interfacial oxide.) A large hole generation is observed in the channel close to drain, due to band-to-band tunneling. (b) A 2D plot showing the hole concentration of the entire device. A large hole concentration on the order of 10^{20} cm^{-3} is observed at the cross section where (a) is taken. 72
6.4 Waveform schematics for writing and reading (a) high-V_T (PGM) and (b) low-V_T (ERS) states. The hold operation is defined as the stage where $V_{GS} = 0$ V. The readouts for both states are conducted at $V_{GS} = 1$ V and $V_{DS} = 0.5$ V. The read time is on the order of ns. 74

6.5 Transient evolution of (a) Hole density in the Si channel and (b) Ferroelectric polarization during zero-bias hold with different ρ. Solid curve: high-V_T phase, dashed curve: low-V_T phase. Reduced hole screening causes the polarization to degrade during hold of high-V_T phase, while the polarization during the low-V_T phase remains at a constant value during longer holds. 74
6.6 A demonstration of the device's V_T shift; as the zero-bias hold time T_{H0} increases from 10 µs to 1 s, the PGM V_T (red) steadily decreases from 1.12 V to 1.02 V, while the ERS V_T (black) remains unchanged. 75
6.7 (a) Log drain current during readout (defined as the readout current) after different zero-bias-hold times, $\rho = 2.25 \times 10^5$. Solid: high-V_T , dashed: low-V_T . The low-V_T readout currents for hold times greater than the polarization switching time are identical and remain flat throughout. The plot is zoomed-in to sub-ns for better visualization of the decay. (b) Maximum read margin, defined as the I_D ratio of the quasi-steady-state drain current for the two states in (a), as a function of hold time. The arrows to the right show the time scale dominated by the LK equation, while the arrows to the left indicate the hole-dynamics-dominated zone, particularly during the high-V_T phase. 76
6.8 (a) Log hole density vs. time during readout in the high-V_T phase for different zero-bias-hold times. The hole density reaches quasi-steady state within 0.5 ns. (b) Hole source current during readout of the high-V_T phase. The hole current decreases with increasing hold time due to lower hole concentration, hence diffusion. The quasi-steady-state hole current at $T_{H0} = 100$ ns is slightly lower than $T_{H0} = 10$ µs because of incomplete polarization switching in the ferroelectric at short time scales. 77
6.9 Energy band diagram at the start (dashed) and end (solid) of the high-V_T state readout. The increases as holes (purple) flow into the source, which increases the energy barrier for the injection of source electrons. 77
6.10 (a) Electron source current during readout of the high-V_T phase. The quasi-steady-state electron current increases with increasing zero-bias-hold time. (b) Conduction band energy of a vertical cross section showing the Si/SiO_2/FE stack at the end of high-V_T readout. The electron quasi-Fermi levels at the source are aligned for all the cases. The reduces as the hold time increases because of reduced polarization, which leads to an increased quasi-steady-state electron source current in (a). 78
6.11 Lateral band diagrams with quasi-Fermi levels taken at Y= 0.011 during the 0 V_{GS} hold with applied drain bias. (a) V_{DS} = 0.5 V (b) V_{DS} = 0.8 V, and (c) V_{DS} = 1.0 V. (d) The channel hole concentration as a function of zero-bias hold time, for the 3 different V_D values. Note that the point at which the hole concentration begins to level off varies significantly with the chosen drain voltage. The high hole concentration and associated ΔV_t below 1 µs is related to incomplete polarization switching. 80

List of Tables

6.1 Selected device paramters 71

Acknowledgments

First and foremost, I would like to thank my research advisor, Prof. Sayeef Salahuddin, for all the kind advice and firm supports over the years. I feel honored and fortunate to be a part of the group, and I really appreciate the impact-driven research culture here at Berkeley. I have grown a lot over the past few years, largely thanks to the role models that I am lucky enough to know in person and learn from. Prof. Salahuddin profoundly inspired me with his keen observations, creativity, and integrity as a researcher, which helped me develop the right attitude and methods to approach different problems. I am also grateful to Prof. Chenming Hu who is also a great source of inspiration and has shared numerous insights and encouragements with me. Both the professors are very kind and down-to-earth, which makes my life as a Ph.D. student delightful. I am immensely thankful for all the supports and the guidance, both technical and personal.

I feel deeply indebted to my undergraduate research advisor, Prof. Cheewee Liu, who opened the door of the fascinating field of semiconductor device research for me. Prof. Liu shaped my self-awareness as an engineer, trained me rigorously and patiently when I was a novice to the field, and encouraged me to pursue a Ph.D. overseas. Without the help, I wouldn't have had the luck to join UC Berkeley nor to build the foundation for the subsequent achievements. Besides, I would like to thank Prof. Junqiao Wu and Prof. Ali Javey for serving on my qualification and dissertation committee. The fruitful discussions with them greatly helped me wrap up my research projects. I also appreciate Prof. Jenn-Gwo Hwu, Prof. Samuel Pan, and many others for all the help and advice. I hope to pass on their kindnesses by staying humble and helping others in the same way as they do.

I hope to express my gratitude to all my family members which is beyond the words. My parents and grandparents raised and educated me well, making my childhood full of care and happiness. They are inseparable from my life. Another special acknowledgement goes to my fiancée, Hsin-Yi, for her continued emotional supports. Her company fills my life in the US with joy and love, constantly bringing motivations and optimisms to me.

I thank all my friends and colleagues at Berkeley, Yen-Kai, Li-Chen, Ming-Yen, Tien-Ning, Hsin-Han, Frank, Tim, Zephy, Fang-Chieh, Chen-Yu, Ping-Hung, Hsiou-Yuan, Yow-Gwo, Ava, Wenshen, Adi, Saavan, Jason, Suraj, Nirmaan, Jiuren, Hong, Korok, Justin, Daewoong, Jong-Ho, Juan, Harshit, Pragya, Avirup, Niklas, Sharon, Neelesh, Michael, Sam, Varun, Alice, Jennifer, Chung-Kuan, and many others, for the help and the good memories. I also would like to thank my mentors, Sheng-Ting and Der-Chuan, for their selfless help and guidance in the early stage of my career. I thank my relatives in the US, Judy, M.-P., Rebecca, Sylvain, and others, for their warmth and hospitality. There are many other people to whom I owe a great deal of thanks. Meeting each of you is a fortune for me, and I wish all the best for you.

Chapter 1

Introduction

1.1 Needs For Electrical Scaling of CMOS Gate Stack

The advance of Complementary-Metal-Oxide-Semiconductor (CMOS) technology has been mainly motivated by the improvement in the power, performance, and area for digital circuits. Several key benchmarks at the MOS Field Effect Transistor (MOSFET) operation level include

- the dynamic energy dissipation for a switching event CV_{DD}^2,
- the intrinsic delay for a switching event CV_{DD}/I_{eff}, and
- the static leakage power for a stand-by device $V_{DD}I_{off}$.

where the C is the effective load per stage including the gate and the fringing field capacitance, V_{DD} the CMOS supply voltage, I_{eff} the effective switching current of the transistor in the ON state, and I_{off} the stand-by current when the transistor is in OFF state. A low C, a low V_{DD}, a low I_{off}, and a large I_{ON} are desired in order to simultaneously achieve low power dissipations and high speed. The last three requirements imply that a sharp transition from OFF to ON state is needed. That is, given a small change in the MOSFET gate-to-source voltage (V_{GS}), the increase of the MOSFET drain current (I_D) should be as large as possible.

For the transition rate of MOSFET in strong inversion regime, commonly characterized by the transconductance ($g_m \equiv \partial I_D/\partial V_{GS}$), a high carrier velocity and a large intrinsic gate capacitance are the key boosters. For the sub-threshold regime where the carrier screening effect is negligible, the transition rate is dominated by the electrostatic control. The corresponding figure of merit is the sub-threshold swing (SS)

$$SS \equiv \frac{\partial V_{GS}}{\partial \log_{10} I_D} = \frac{\partial V_{GS}}{\partial \phi_S}\frac{\partial \phi_S}{\partial \log_{10} I_D} = \frac{\partial V_{GS}}{\partial \phi_S}\frac{kT}{q}\ln(10) \tag{1.1}$$

where k is the Boltzmann constant, T is the operating temperature, q is the charge of an electron, and ϕ_S is the surface potential of the channel that can be determined by the

capacitive voltage divider.

$$\frac{\partial V_{GS}}{\partial \phi_S} = 1 + \frac{C_{Si}}{C_g} \quad (1.2)$$

Here the C_g is the electrostatic capacitance between gate and the top of the potential barrier (TOB) in the silicon channel, and the C_{Si} is the capacitance of the TOB that is ascribed to the silicon body. When the mobile carrier has negligible effects on the electrostatics, C_{Si} includes the coupling from the source, the drain, the substrate. Innovative device structures such as the fully-depleted silicon-on-insulator (FDSOI) MOSFETs, FinFETs, and gate-all-around (GAA) MOSFETs are effective strategies to reduce the C_{Si} and hence suppress the short channel effects [1]. However, further scaling of the gate length (L_g) inevitably increases the finite C_{Si}. Besides, available high-mobility channel materials such as germanium has a larger permittivity than silicon, which can also worsen the C_{Si}. A large C_g is still required to maintain a good gate control. Therefore, the intrinsic gate capacitance is important for achieving sharp transitions in terms of both g_m and SS. The high-κ-metal-gate process has been historically one of the most important solutions in this regard, and it is adopted in the state-of-the-art CMOS technology. Note that, although it may seem unfavorable that the intrinsic gate capacitance increases the C that is proportional to the dynamic energy dissipation and intrinsic delay, the corresponding enhancement in I_{eff} and the allowed reduction in V_{DD} allow for improvements in the two metrics given that the parasitic capacitance constitutes a large proportion of C in modern technologies. As a result, a high capacitance of the gate stack is generally desired and has become increasingly important for MOSFET scaling.

Specifically, the International Roadmap For Devices and Systems (IRDS) [2] has projected a reduction of V_{DD} from the 0.75V as of the year of 2021 to 0.55V by 2034, while a similar ON-OFF ratio to the current technology needs to be maintained. One of the major challenges for this goal is in the gate stack. The roadmap calls for an inversion layer thickness (t_{inv}), which is the effective SiO_2 thickness (EOT) for the inversion capacitance C_{inv}, of 0.90nm.

$$t_{inv} = \frac{\epsilon_{SiO_2}}{C_{inv}} = \frac{3.9\epsilon_0}{C_{inv}} \quad (1.3)$$

where ϵ_{SiO2} and ϵ_0 are the perimittivty of SiO_2 and vacuum, respectively. However, this has yet to be reliably achieved through classical high-κ gate stack. For a high-κ-metal-gate stack in strong inversion.

$$t_{inv} = EOT_{cl} + EOT_{IL} + EOT_{HK} \quad (1.4)$$

where EOT_{cl} is the EOT for the semiconductor inversion layer of which the charge centroid is located at a finite distance away from the semiconductor/insulator interface, EOT_{IL} is the EOT of the interfacial layer (IL) that is between the semiconductor and the high-κ layer, and EOT_{HK} is the EOT of the high-κ layer. The last two terms adds up to the EOT that is purely ascribed to the gate stack ($EOT_{stack} = EOL_{IL} + EOT_{HK}$, also commonly referred to as simply "EOT" in the literature.) Calculations considering quantum confinement effects [3] indicate that the EOT_{cl} is approximately 0.4nm and 0.3nm for pure silicon and pure

germanium channels, respectively, when the sheet electron density is $1 \times 10^{13} cm^{-2}$ which is a typical value for the ON state. Since direct contact of high-κ materials on silicon would lead to an unstable interface, an SiO_2 IL is needed to passivate the silicon surface before a the high-κ layer is deposited [4, 5]. Typically, an SiO_2 IL thickness of 0.7nm or above can be engineered with good quality [4]. While EOT_{stack} can be reduced by further reducing the IL through processes such as SiO_2 scavanging, it has a drastic impact on the carrier mobility, gate leakage, and reliability for sub-monolayer (<0.5nm) IL [5, 6, 7, 8]. The limitations of EOT_{cl} and EOT_{IL} leave a room that is less than 0.1nm for EOT_{HK}, of which a mature solution has not yet existed. Innovations are needed to enable good gate control for future CMOS technology nodes.

1.2 Ferroelectric Negative Capacitance

The seminal paper [9] in 2008 has triggered an extensive research interest in the negative capacitance (NC) effect that can stem from the polarization response to an external electric field of ferroelectric (FE) materials. The negative capacitance effect has enormous implications on transistor scaling because it would imply a "potential amplification" effect that can significantly boost the gate capacitance despite a large physical gate stack thickness. It can also enable classically impossible phenomena that will be discussed in 1.3 and throughout the dissertation.

Negative Capacitance State of an FE Capacitor

Based on the Landau-Devonshire model for ferroelectricity, the Gibbs free energy density for a ferroelectric (FE) material system with 1-D macroscopic polarization and electric field can be formulated as [10]

$$\mathcal{G}_{FE} = \alpha P_{FE}^2 + \beta P_{FE}^4 - P_{FE} \times E_{FE} + h.o.t. \tag{1.5}$$

where P_{FE}, the order parameter, is the FE polarization, and E_{FE} is the net electric field. Only terms relevant to P_{FE} are written here, while a more complete form including electric field energy will be specified in the next subsection. Here the lowest-order energy coefficient α is negative. For a standalone system where $E_{FE} = 0$, the state $P_{FE} = 0$ would correspond to a local free energy maximum. Therefore, it represents an unstable state that is not energetically favored, and the thermal equilibirium is reached instead at other states with large $|P_{FE}|$ (the remnant polarization) that correspond to the free energy minima due to the higher-order energy terms. For a steady state, $\partial \mathcal{G}_{FE} / \partial P_{FE} |_{E_{FE}} = 0$, which results in the relationship (Fig. 1.1(a)(b))

$$E_{FE} = 2\alpha P_{FE} + 4\beta P_{FE}^3 + h.o.t. \tag{1.6}$$

The displacement D as defined in the Maxwell equations equals to

$$D = \epsilon_b E_{FE} + P_{FE} \tag{1.7}$$

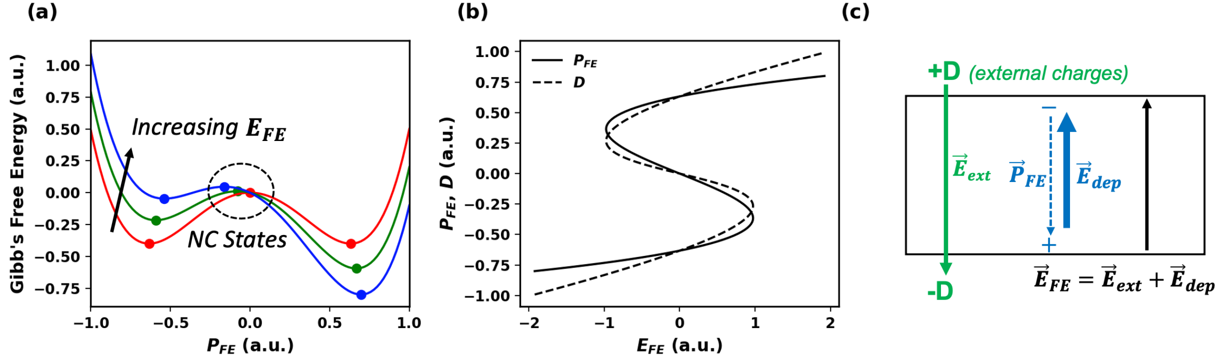


Figure 1.1: Schematic (a) $\mathcal{G}_{FE} - P_{FE}$ relationships at different E_{FE}. Possible steady states are marked. For the NC states, P_{FE} decreases as E_{FE} increases in contrast to classical trend. However, they are not stable in a bulk system because they correspond to local maxima of $\mathcal{G}_{FE}$. (b) $E_{FE} - P_{FE}$ (solid) and $D - P_{FE}$ (dashed) relationships. Both have negative slopes for small $|P_{FE}|$, corresponding to the NC state. (c) Illustration of different fields for an NC state. P_{FE} responds in the same direction as E_{ext}. However, the depolarization field generated by P_{FE} ($E_{dep} = -P_{FE}/\epsilon_b$) is stronger than E_{ext}, reversing the direction of the net electric field E_{FE}.

where $\epsilon_b \geq \epsilon_0$ is the background permittivity that has been generalized to include possible DE responses that takes place in the system in parallel with the FE response. The capacitance density $\mathcal{C}_{FE}$ is then calculated as

$$\mathcal{C}_{FE} = \frac{\partial D}{\partial E_{FE}} = \epsilon_b + \frac{1}{2\alpha + 12\beta P_{FE}^2 + h.o.t.} \tag{1.8}$$

If able to be stabilized, the small-P_{FE} state could have $\mathcal{C}_{FE} < 0$ because of the negativity of α.

Note that, although P_{FE} and E_{FE} are in the opposite directions for the NC state, P_{FE} responds in the same direction as the the external electric field (E_{ext}) that stems from charges that are external to the system (Fig. 1.1(c).) Since the charge density on the electrodes are determined by the displacement in the capacitor, when the depolarization field due to DE response is included in E_{ext}, we have

$$E_{ext} = \frac{D}{\epsilon_b} = E_{FE} + \frac{P_{FE}}{\epsilon_b} \tag{1.9}$$

E_{ext} and P_{FE} always have the same sign. The net electric field can also be expressed as $E_{FE} = E_{ext} - P_{FE}/\epsilon_b$, which illustrates how the NC effect arises from the P_{FE} response. The depolarization field due to the polarization ($E_{dep} = -P_{FE}/\epsilon_b$) exceeds the E_{ext} that drives the polarization, resulting in the net electric field E_{FE} that is in the opposite direction as D and E_{ext}.‘

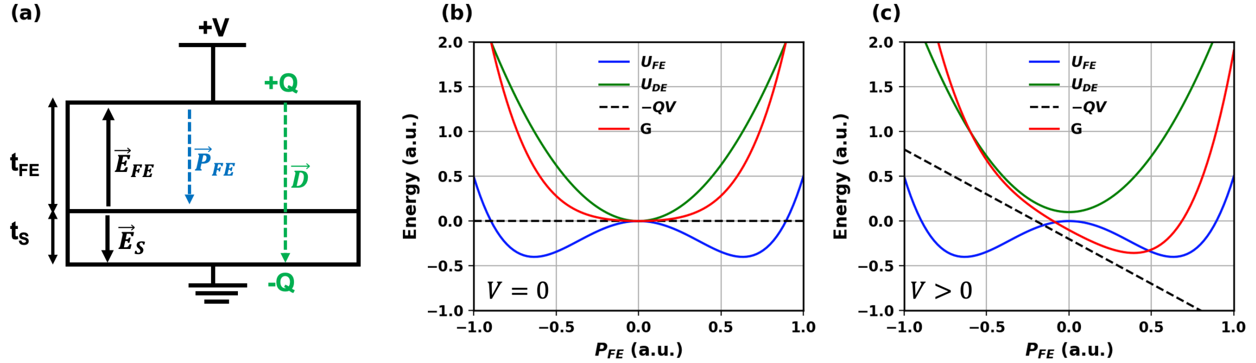


Figure 1.2: (a) Illustration of the fields in the FE-DE stack with applied voltage when FE is in the NC state. The displacement D is continuous throughout the layers, which connects E_S and E_{FE} to V and P_{FE}. (b)(c) The energy landscapes of a FE-DE stack system under different applied voltages. The system satisfies condition 1.20, so G is stabilized at the low-P_{FE} state, and the hysteresis is eliminated. The overall capacitance of the system is positive, so as V increases, Q and D increases, which drives P_{FE} in the same direction, while E_{FE} moves in the opposite direction due to the NC effect.

Energy Landscape of an FE-DE Stack

The following arguments are based on [9] but have been extended with derivations to clarify the global energy landscape in order to resolve confusions that sometimes occurred in the research community. The authors argue that while impossible in a standalone system, the stabilization of small-P_{FE} can be achieved in terms of the global energy picture through series connections of an FE capacitor to a normal dielectric (DE) capacitor. Consider two parallel capacitors connected in series (Fig. 1.2(a)), if the two capacitors have the same area (A), the total free energy G of the series capacitors with applied voltage V is [11]

$$G = U_{FE} + U_{DE} - QV \tag{1.10}$$

$$U_{FE} = At_{FE}(\alpha P_{FE}^2 + \beta P_{FE}^4 + h.o.t.) \tag{1.11}$$

$$U_{DE} = A\left(t_{FE}\frac{\epsilon_b E_{FE}^2}{2} + t_S\frac{\epsilon_S E_S^2}{2}\right) \tag{1.12}$$

where t_S and t_{FE} are respectively the thicknesses of the DE and FE capacitors, E_S is the electric field in the DE capacitor, ϵ_S is the permittivity of the DE material, and Q is the total charge on the electrode. Here U_{FE} is the Helmholtz free energy associated with the FE polarization response, U_{DE} is the Helmholtz free energy associated with the electric field energy and the DE responses in both the FE and DE layer, and QV is the Legendre transformation term to formulate Gibb's free energy with a fixed applied voltage.

When normalized to area, we have

$$\frac{G}{A} = t_{FE}(\alpha P_{FE}^2 + \beta P_{FE}^4 + ... + \frac{\epsilon_b E_{FE}^2}{2}) + t_S \frac{\epsilon_S E_S^2}{2} - DV \tag{1.13}$$

where D is the displacement that is uniform throughout the capacitors. It may seem like E_S is a new degree of freedom in eq. 1.13 due to introduction of the DE subsystem, but it is not. Both E_S and E_{FE} are functions of V and P_{FE} as dictated by the Maxwell equations. According to the continuity of displacements, we have

$$D = P_{FE} + \epsilon_b E_{FE} = \epsilon_S E_S \tag{1.14}$$

Because the system is non-magnetic, we have (Fig. 1.2(a))

$$V = t_{FE} E_{FE} + t_S E_S \tag{1.15}$$

Solving the above system of equations, we have

$$E_{FE} = \frac{\epsilon_S V - t_S P_{FE}}{\epsilon_S t_{FE} + \epsilon_b t_S} \tag{1.16}$$

$$E_S = \frac{\epsilon_b V + t_{FE} P_{FE}}{\epsilon_S t_{FE} + \epsilon_b t_S} \tag{1.17}$$

Therefore, the free energy in eq. 1.13 can be expressed as a function of P_{FE} and V, analogous to the form of eq. 1.5. Given an applied voltage, the energy landscape of the entire system can be aligned to P_{FE}. We then examine the first-order and second-order properties of the P_{FE}-G relationship.

$$\left.\frac{\partial(G/A)}{\partial P_{FE}}\right|_V = t_{FE}(2\alpha P_{FE} + 4\beta P_{FE}^3 + h.o.t.) + \frac{P_{FE}}{\epsilon_S/t_S + \epsilon_b/t_{FE}} - \frac{V}{1 + \frac{t_S}{\epsilon_S}\frac{\epsilon_b}{t_{FE}}} \tag{1.18}$$

$$\left.\frac{\partial^2(G/A)}{(\partial P_{FE})^2}\right|_V = t_{FE}(2\alpha + 12\beta P_{FE}^2 + h.o.t.) + \frac{1}{\epsilon_S/t_S + \epsilon_b/t_{FE}} \tag{1.19}$$

The last two terms in the RHS of eq. 1.18 can be rearragned according to eqs. 1.14 and 1.15

$$\begin{aligned}
\frac{P_{FE}}{\epsilon_S/t_S + \epsilon_b/t_{FE}} - \frac{V}{1 + \frac{t_S}{\epsilon_S}\frac{\epsilon_b}{t_{FE}}} &= \frac{1}{\epsilon_S/t_S + \epsilon_b/t_{FE}}\left[P_{FE} - \frac{\epsilon_S}{t_S} t_{FE} E_{FE} - \epsilon_S E_S\right] \\
&= \frac{1}{\epsilon_S/t_S + \epsilon_b/t_{FE}}\left[P_{FE} - \frac{\epsilon_S}{t_S} t_{FE} E_{FE} - D\right] \\
&= \frac{1}{\epsilon_S/t_S + \epsilon_b/t_{FE}}\left[-\frac{\epsilon_S}{t_S} t_{FE} E_{FE} - \epsilon_b E_{FE}\right] \\
&= -t_{FE} E_{FE}
\end{aligned}$$

Plugging this back into eq. 1.18 leads to the same P_{FE}-E_{FE} relationship as eq. 1.6. This confirms that in the combined system the FE layer still follows the S-curve as dictated by

the Landau energy landscape, and the NC state is a possible solution. When $t_S = 0$, the second term in the RHS of eq. 1.19 vanishes, leading to the same condition as the standalone capacitor. However, with increasing t_S, the second term increases, making the overall second-order derivative positive even when the first term in eq. 1.19 is negative for the small-P_{FE} NC state. It can be easily shown from eq. 1.19 that, if the FE capacitance density $\mathcal{C}_{FE}$ (whose unit equals to perimittivity) is negative, the stability condition follows

$$\left.\frac{\partial^2 G}{(\partial P_{FE})^2}\right|_V > 0 \iff \frac{\epsilon_S}{t_S} < -\frac{1}{t_{FE}}\left(\frac{1}{2\alpha + 12\beta P_{FE}^2 + h.o.t.} + \epsilon_b\right) = \frac{|\mathcal{C}_{FE}|}{t_{FE}} \tag{1.20}$$

Equivalently, the FE negative-capacitance state becomes stable if the DE series capacitance is low enough to make the overall capacitance of the system ($C_{tot} = (t_S/\epsilon_S + t_{FE}/\mathcal{C}_{FE})^{-1}$) positive.

The energy landscape provides a physical picture as to how the NC stabilization takes place (Fig. 1.2(b)(c).) The DE series connection penalizes the large polarization state, and hence the stabilized low-P_{FE} state becomes the only stable state of the overall system. The system becomes hysteresis-free and has a larger C_{tot} than the DE-only capacitor (ϵ_S/t_S) when P_{FE} is small enough so that $\mathcal{C}_{FE} < 0$.

Microscopic Explanations to Negative Capacitance Effects

Complementing the physical picture with free energy, [12] provides a force-based explanations to the S-shaped P_{FE}-E_{FE} relationship and the stabilization of the NC state. The authors considered uniform atomic displacements withing a ABO_3 perovskite structure, a classical class of FE materials. The is polarization and net electric field is defined to be in the z direction. The polarization P_{FE} is dominated by and is proportional to the displacement of the the B cation atoms (z_B.) z_B is determined by the force balance on the B cations $F_{loc}^B = q_B E_{loc}^B$, where the cation charge (q_B) is fixed, and thus the force is proportional to the local electric field on the atom E_{loc}^B.

$$E_{loc}^B = E_{int}^B + E_{ext} \tag{1.21}$$

where E_{int}^B is the electric field due to other charges in the FE material, and E_{ext} is the external electric field following eq. 1.9. Note that the E_{FE} and E_{ext} are macroscopic electric fields, which differs from the microscopic local electric field E_{loc}^B. The authors of [12] found that when $|z_B|$ is small, the internal electric field due to the O anoins would further polarize B in the already polarized direction. After rearrangement of terms, the lowest order terms for E_{loc}^B can be expressed in terms of the macroscopic fields

$$E_{loc}^B = -2\alpha P_{FE} + h.o.t. + E_{FE} \tag{1.22}$$

where α is negative. $E_{loc}^B = 0$ for a steady state to be reached. This corresponds to $E_{FE} = 2\alpha P_{FE} + h.o.t.$, which coincides with the first-order condition for the free energies from the previous subsection.

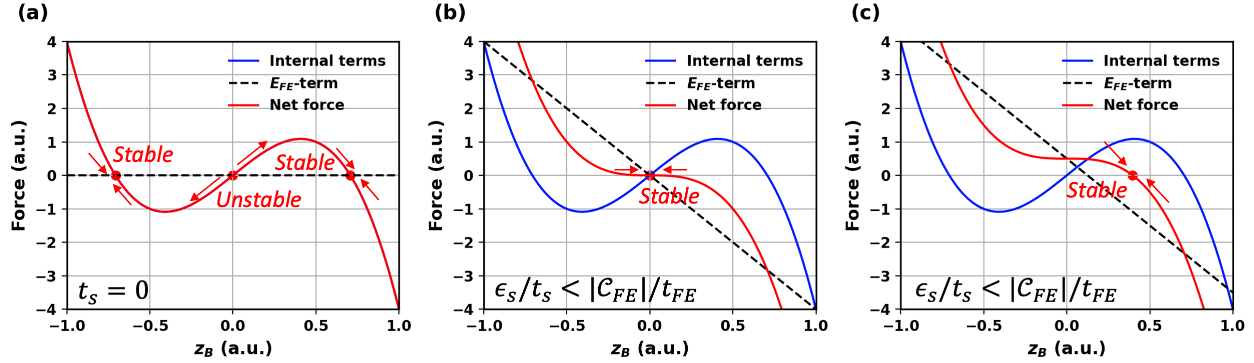


Figure 1.3: Relationship between the cation displacement and the forces on the cation atom. The net force is decomposed into internal terms that are unrelated to the applied field and the E_{FE} term. (a) is without a DE layer and V=0, so $E_{FE} = 0$ regardless of z_B and P_{FE}. The internal terms make the NC state unstable, while only the remnant polarization states are stable. (b) and (c) are with a DE layer that is thick enough to stabilize the NC state which becomes the only steady state. The FE polarization induces a depolarization field that restores the cation from fluctuations. When the voltage bias is changed from (b) to (c), the dashed curve shifts vertically, which results in a new steady state.

The stability of a polarization state can be intuitively understood with the force picture. For a standalone FE capacitor, E_{FE} is fixed, so if P_{FE} fluctuates slightly from the steady state, the net force F^B_{loc} is in the same direction as the deviation (Fig. 1.3(a),) which would further polarize the P_{FE}, i.e. a positive feedback is formed. Therefore, the NC state $P_{FE} \approx E_{FE}/2\alpha$ is unstable. However, the E_{FE} becomes a function of P_{FE} when a DE capacitor is connected in series (eq. 1.16,) so eq. 1.22 becomes

$$E^B_{loc} = -2\alpha P_{FE} + h.o.t. + \frac{\epsilon_S V - t_S P_{FE}}{\epsilon_S t_{FE} + \epsilon_b t_S} \tag{1.23}$$

In terms of the changes, we have

$$\Delta E^B_{loc} \approx \left(-2\alpha P_{FE} + \frac{-t_S}{\epsilon_S t_{FE} + \epsilon_b t_S} \right) \Delta P_{FE} \tag{1.24}$$

Although the first term in the parenthesis is positive and tends to further polarize the material, the last term is in the opposite direction to ΔP_{FE} (Fig. 1.3(b)(c).) Effectively, a “depolarization” field is introduced due to the imperfect screening of the polarization change. The NC state is stabilized if a negative feedback on the P_{FE} can be formed. This is achieved if the coefficient for ΔP_{FE} in the RHS of eq. 1.24 is negative, which corresponds to $\epsilon_S/t_S < -\left[(2\alpha + 12\beta P^2_{FE} + h.o.t.)^{-1} + \epsilon_b\right]/t_{FE}$, identical to the conclusion derived from the energy landscape in the previous subsection.

Multi-Domain Considerations

The above analyses are based on the idealization of a uniform electric field and polarization throughout the material systems. In reality, domains with different polarization states may exist within the same FE layer. [13] shows with simplistic two-domain models that, if there is a metallic inter-layer between the FE and the DE layer, charge smoothing effect due to the metal would screen out the depolarization field that may stem from the DE. In this case, domains with opposite remnant polarizations are energetically preferred over the NC state. The result suggests that direct interfacing the FE and DE layers are required to establish the continuous boundary condition for electrical displacements.

Note that stable NC effect can also take place in the transition region between domains (the "domain wall") where the P_{FE} is close to zero [14, 15, 16]. This was experimentally demonstrated on the samples with polarization vortices in [17]. The "domain softening" effect on the FE domain structure due to the direct contact to a DE layer has been theoretically predicted before NC effect gets extensive attention [14]. Calculations with continuous spatial variations [18, 16] show that there also exists a critical thickness that is associated with the domain wall energy for direct DE-FE stack to maintain single-domain ferroelectricity and to maximize the NC effect. Consistent with the above discussions, single-crystalline FE experiments with externally-connected DE-FE systems [19] or thick FE-DE stacks [20] are characterized with hysteretic behaviours, despite the measured potential amplification effect and the expected stabilized condition. On the other hand, with carefully engineered DE-FE stack, the potential amplification effect without hysteresis has been measured [21, 22].

Practically, domain formations would imply possible hysteresis [23] and non-homogeneous potential that is undesirable for CMOS logic devices. Besides, the overall amplification effect would be significantly degraded in this case [24]. Therefore, for integration of quasi-static NC effect onto CMOS, we need

- direct deposition of FE layer in the front-end gate stack. There should be no metal inter-layer between the FE layer and the channel.
- FE layer should be thin enough for NC stabilization and to suppress formation of hard domains.

Fortunately, this can be achieved with novel materials system as we will discuss in 1.3.

TCAD Simulations

Recent development of commercial technology computer-aid design (TCAD) [25] tool has incorporated the Landau-Kalatnikov (LK) equation into the device simulator. At a given position,

$$\rho \frac{\partial P_{FE}}{\partial t} + \nabla_{P_{FE}} U = 0 \tag{1.25}$$

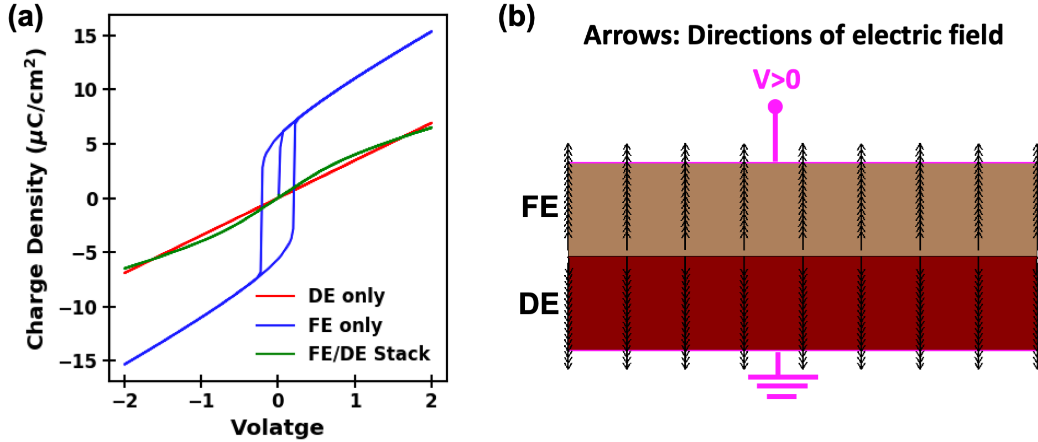


Figure 1.4: (a) Simulated Q-V results for a DE capacitor, an FE capacitor, and a stack structure combining both. LK model is used to simulate the P_{FE} responses. (b) A snapshot of the electric field profile for the FE-DE stack structure in NC state.

where ρ is a viscosity associated with the polarization switching dynamics, and U is the free energy.

$$U = \alpha P_{FE}^2 + \beta P_{FE}^4 + \gamma P_{FE}^6 + g|\nabla P_{FE}|^2 - E \cdot P_{FE} \tag{1.26}$$

which includes the Landau energy expansion up to the sixth order (with coefficients α, β, and γ) and a polarization gradient term (with coupling coefficient g). Partial differential equations formulated with LK equation, the non-linear Poisson equation for electrostatics, and the continuity equations for mobile carrier transports can be solved self-consistently, which enables modeling of the physical details for semiconductor devices fabricated with ferroelectric materials. The interplay between the P_{FE} and the electrostatics of short-channel transistors is one of the major focuses of this dissertation.

We demonstrate the capability of the LK model through simplistic phase field simulations with FE and FE-DE stack structures. Transient analysis are performed by coupling eq. 1.25 and the Poisson equation ($\nabla D = 0$ for insulators without fixed charges.) The results are shown in Fig. 1.4(a). The capacitors are applied with triangular-wave voltage signal, and the resultant charge-voltage (Q-V) relationship shows significant hysteresis if only an FE capacitor is simulated. However, for a stabilized FE-DE structure, the simulated Q-V is hysteresis-free, while a capacitance boost manifesting as the increased slope as compared to the DE-only case is observed in the middle. The capacitance boost diminishes for large Q where the high-order terms in U dominate the P_{FE} response. With an appropriate system configuration, single FE domain can be achieved in the stack structure, and the net electric field are in the opposite directions for the FE and DE layer when the system is in the NC state (Fig. 1.4(b).)

1.3 Negative Capacitance Transistors

Negative capacitance field-effect transistors (NCFET) relies on the NC effect of the ferroelectric gate insulator to induce differential amplification of the gate signal. The enhanced gate control thus provides favorable CMOS scaling prospects [9]. NCFET research has been significantly facilitated by the discovery of the ferroelectric orthorhombic phase in the Hf-based materials system [26]. The two key aspects of Hf-based ferroelectrics make them promising candidates for CMOS integration.

- Process and materials compatibility [27]. Hf-based ferroelectrics can be fabricated through atomic-layer deposition (ALD) that has been deployed for high-κ metal gate process. HfO_2 is the standard high-κ material in novel CMOS gate stacks, so relatively few modifications to the mature process is needed for the integration of Hf-based FE gate stack. Integration of Si-doped HfO2 ferroelectrics into the front end of 14nm FinFET process has been demonstrated [28].

- Low critical thickness. Ferroelectricity has been demonstrated for Hafnium-Zirconium-Oxide (HZO) films under 2nm [29], matching the gate stack thickness requirements for 3-D MOSFETs.

Substantially improved I-V characteristics and non-classical sub-threshold behavior have been observed for NCFETs with HZO gate stack [30, 31, 32, 33, 34, 35] that has a polar (ferroelectric or antiferroelectric) order [36, 37, 38].

One of the prospects of NCFET is to make C_g in eq. 1.2 negative, which can lead to a sub-thermal sub-threshold swing (SS), a negative Drain-Induced Barrier Lowering (DIBL), or a negative differential resistance [39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49]. However, C_{Si} is a strong function of the channel mobile charge, which makes it difficult to satisfy the stabilization condition $(1/C_{Si} + 1/C_g)^{-1}$ in the strong inversion region where C_{Si} is large. Unless C_g is also highly non-linear, a sub-60mV/dec SS achieved through negative C_g would usually correspond to some degree of hysteresis [50]. Nevertheless, the benefit of the NC gate stack is not restricted to apparently non-classical I-V characteristics. The NC effect also opens an opportunity to reduce the gate stack EOT without scavenging the IL, which is an already significant improvement that is sought after as discussed in section 1.1. EOT_{stack} lower than the IL thickness can be achieved with NC gate stack, while a chemically and electrically stable IL maintains good mobility. Projections on the high-speed operation of NCFET based on the experimental results will be discussed in chapter 5.

Negative Capacitance Switching Dynamics

Whether the negative capacitance response can be fast enough to support high-speed circuit operations was another key aspect affecting the applicability of NCFET. Some research works [51, 52, 53] aimed to answer this question by inferring the LK switching dissipation dynamics

through electrical characterizations of MFM (Metal-FE-Metal stack) or MFMIM (Metal-FE-Metal-DE-Metal) capacitors. However, the quasi-static NC effect is not expected to be established in this class of systems as discussed in section 1.2 and experimentally confirmed in [54], and the polarization responses will be largely dominated by the nucleation-limited switching dynamics [55, 56] instead of the small-signal NC effect. Besides, some technical limits in the capacitor characterization methods make it difficult to characterize the response below the nanosecond regime that is however relevant to the high-speed circuit operations. A relatively large capacitor area is required for the capacitive response signal to be detected, which makes the measurement susceptible to parasitic resistance in the laboratory setup. For example, the ρ value extracted in [51] would correspond to about 30Ω lumped resistance for the whole system, which may include pessimism due to parasitic resistance that is of the same order. The effect can be mitigated by reducing the capacitor size and by carefully de-embedding the parasitics, and a much reduced polarization switching time below 1ns [55] was reported accordingly, but accurate ρ extraction has yet to be reported through this characterization method. Besides, as we will see in the following, the LK equation is insufficient to describe the polarization dynamics in ultra-fast regime.

The authors of [57] revisited the theory of FE switching dynamics with Drude oscillator model. The analysis reveals the importance of an inertia term that is associated the second-order time derivative of P_{FE}, and the LK equation is a low-speed approximation to the complete dynamics which is valid when the oscillation frequency is low enough so that the $\partial^2 P_{FE}/(\partial t)^2$ can be neglected. The formalism enables extractions of the dynamical parameters through infrared absorption measurements, and the authors estimated the NC response time for logic device switching to be lower than 0.5ps [57]. Ring-oscillators consisting of NCFETs reported in [28] are systematically analyzed in [58] to verify the prediction. Details of the study will be discussed in chapter 2.

Anomalies of NCFET Electrical Characteristics

Sub-60 mV/decade operation and negative DIBL have been the major characteristics that are used to identify negative capacitance operation in experimental data. It has not been fully established whether this can be robustly achieved through quasi-static NC effect because of the difficulties discussed above. However, other interesting aspects of the potential amplification effect due to the stabilized "S"-shaped P_{FE}-E_{FE} response can also lead to characteristics that can not be explained with conventional dielectric theory. In particular, we observed anomalously high improvements in the gate control for short channel devices in terms of the gate-length scaling trend [59], and the behaviours are studied in chapter 3. On another NCFET wafer [35], an unconventional SS-I_D trend as opposed to the mobile charge screening effect was observed [60], and this is analyzed in chapter 4.

To understand the origin of these behaviors, controlled experiments with conventional gate stacks are first characterized to gather an understanding to the channel electrostatics and hence the expectations for conventional EOT scaling (Fig. 1.5.) The discrepancies between the expectations and the observed NCFET characteristics are then ascribed to

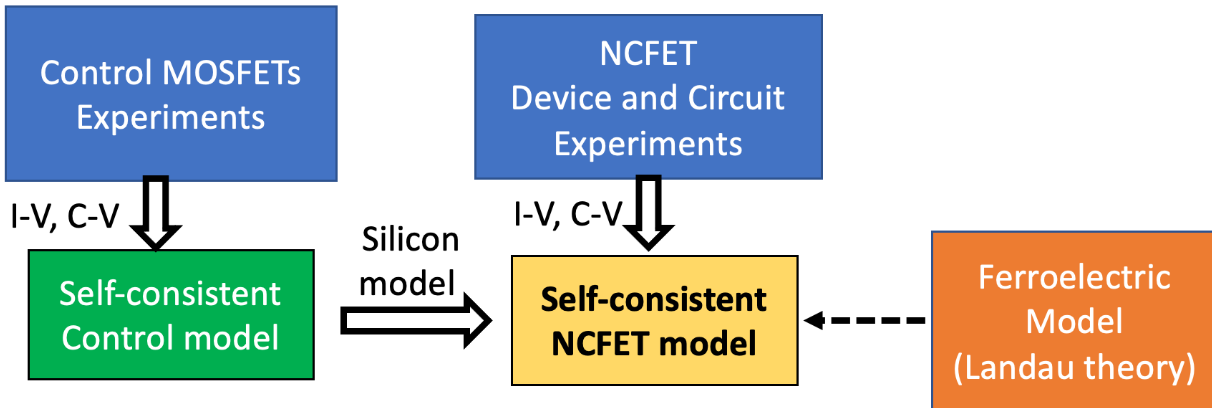


Figure 1.5: NCFET modeling methodology flow with controlled experiments.

the introduction of the NC gate stack. Self-consistent NCFET simulations through TCAD provides a systematic way to combine the understanding of the silicon channel and the Landau phenomenological model which captures the NC effect. Through the method, it was found in the studies [59, 60] that the anomalous characteristics can be explained by the non-linear potential amplification effects, and the NCFET experimental results are successfully reproduced by the TCAD models.

1.4 FeFET Memory

When the FE layer in the MOSFET gate stack is thick enough to overcome the depolarization field, the remnant polarization states of the FE layer modulate the built-in electric field in the gate stack and hence the MOSFET threshold voltage. Ferroelectric field-effect transistors (FeFET) utilize this effect to realize non-volatile memory states. FeFETs have become an area of intense research focus due to their compact cell size and the potential for low-power nonvolatile memories [61]. Experimental designs, such as a 2T FeFET-based nonvolatile memory with separate read and write paths, have been proposed [62]. Furthermore, the potential of FeFETs is expanding, as once-common shortcomings in forming effective 1T-FeFET memory arrays, such as low endurance, are overcome through experimentation in design of material parameters and structure [63, 64, 65, 66]. Very high endurance exceeding 10^{10} cycles have been recently demonstrated on FDSOI wafers [67], which bodes well for future in-memory computing applications. In chapter 6, the interplay between the excess minority carriers for fully-depleted channel and the FE polarization will be analyzed, and its implications on the FeFET design will be discussed.

Chapter 2

Response Speed of Negative Capacitance FinFETs

2.1 Introduction

One question was often raised as to the limit of speed of NC effect, especially regarding whether it hampers the application to state-of-art CMOS logic of which the switching time is of the order of picoseconds. Although theoretical analysis [57] has shown that the intrinsic response time is expected to be $<$0.5 ps, it is critical to validate the prediction on a ferroeletric-dielectric thin-film stack, preferably on transistors fabricated with the process flows of interest. In general, Ring Oscillators (ROs) fabricated with advanced technology provide an appropriate platform to study the transient response of the CMOS transistor and gate stack because of its relevance to real-world application and its well-engineered properties.

- MOSFETs are charge sensors with high sensitivity. The NC switching response is reflected in the channel charge and hence the drain current, which produces robust signal that is easier to measure than the gate impedance.

- The RO operates with large signal ranging from the source rail voltage (VSS) to the drain rail voltage (VDD) at the MOSFET gate node. Therefore, the response of gate stack can be probed at a wide voltage range, and fine-grained information can be extracted through measurements with varied supply voltages.

- The RO operates at a speed that best characterizes novel technologies. Typically, the fan-out 3 (FO3) or fan-out 4 (FO4) structure are representative of the stage concatenation in logic circuit design. As we will see in the following sections, the corresponding stage delay is lower than 10 ps for the NCFET RO in this study, which is in the relevant regime for circuit operation in foreseeable technology nodes [2].

- The parasitic resistance in terms of gate stack charging is minimized throughout the strong-inversion region of the MOSFET through extensive channel, source and drain, and contact optimizations. This mitigates the current starvation problem that often appears in laboratory setup with large capacitors.

In this context, we study the fabricated ROs at high speed. The postulate is that if the NC gate stack cannot respond to these frequencies, one should observe an extra slowdown beyond the usual circuit simulation. That is, the quasi-static I-V and C-V characterized at DC would not be applicable to modeling high-speed operation in the $\sim 10^{11}$Hz regime, and hence the RO characteristics would not be captured by SPICE simulations with the quasi-static models. On the flip side, quasi-static behaviours at high frequency can be validated if RO operations are successfully captured by the simulations without extra assumptions of non-quasi-static (NQS) behaviors, indicating no slowdown for the NC effect introduced.

To test the hypothesis of fast NC response, we have first developed a compact model based on the BSIM CMG framework that fits the DC transfer and output characteristics of the NCFETs. Next, this compact model is fed into a RO simulator to extract the capacitances in the RO by matching the measured speed and active current (IDDA) drawn by the RO. Next, the RO characteristics reaching up to delay¡20ns are compared with simulation. We find that the DC models accurately predicts these parameters, thereby demonstrating that the circuit behavior at these high frequencies is reflective of the DC characteristics. This in turn indicates that there is no discernible delay that can be attributed to the speed limit of the NC effect down to the shortest measured delay, which in our case, was 7.2 ps. [1]

2.2 Validation Methodology

In this section, we discuss the effect of supply voltages on the circuit statics and transients. We then discuss how the trends in experimentally measurable RO characteristics, namely the quiescent current (IDDQ), IDDA, and the effective capacitance (C_{eff}), reflect the MOSFET device characteristics. This provides a basis for the combined modeling-experimental study.

We start by defining the metrics. The RO test structure consists of an FO3 inverter chain with 100 stages and a feedback connecting NAND-2 stage of which one input is the enable pin (Fig. 2.1). IDDQ is characterized with the enable pin voltage set to logic 0 (V_{SS}, also defined as the reference voltage 0V), while IDDA is characterized with enable pin set to logic 1 (V_{DD}), and C_{eff} requires the knowledge of both.

Quiescent current

When the enable pin is set to 0, the output of the NAND gate is locked at logic 1, and the subsequent inverters reach steady state with input and output alternating among logic 0's and 1's. Under this circumstance, the static leakage current for each inverter is dominated by

[1] The content of this chapter is based on [58].

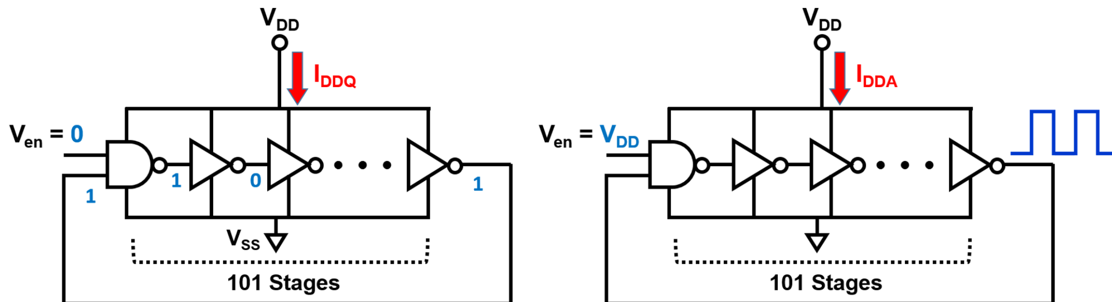


Figure 2.1: Schematics of the 101-stage ring oscillator when it is disabled (left) for IDDQ measurements and when it is enabled (right) for IDDA and delay measurements. Fan-out loads are not shown.

the either the NFET (V_{GS}=0V and V_{DS}=V_{DD}) or the PFET (V_{SG}=0V and V_{SD}=V_{DD}), depending on the bias of the inverter gate. The total leakage current drawn from the supply by the driver inverters (excluding the fan-out loads) is defined as IDDQ. This characterizations serve two purposes in this study.

- Because the DC model is characterized from a standalone test structure other than the RO where no direct probing to single device is available. Checking whether the IDDQ at various supply voltages agree with the DC model provides a reference as to whether the RO devices are consistent with the standalone devices. Particularly, it verifies the threshold voltages, the gate control in sub-threshold regime, and the gate-induced drain leakage effect which reflect the doping profile.

- It offsets the current when the RO circuit is oscillating, and thus needs to be subtracted when the effective switching current is extracted.

Active Current and Delay

When the enable pin is risen to logic 1, the NAND gate becomes logically an inverter for the other input pin, and an switching at the NAND gate output is triggered. The switching events at subsequent inverting stages are then triggered sequentially. As the total number of inverting stages is odd (101) for the ring, the signal triggers a new series of switching events with the opposite edge unateness when it propagates back to the original node, resulting in self-sustaining oscillations. The IDDA is measured in the same way as IDDQ but under the oscillatory operation and is averaged over time.

The active current is directly affected by the transistor conduction current and thus is of central importance for validating the current response at high speed. A well known

relationship for modeling the effective drive current (I_{eff}) of CMOS inverter is [68]

$$I_{eff} = (I_L + I_H)/2 \tag{2.1}$$

where I_L is the drain current at $V_{GS} = V_{DD}/2$ and $V_{DS} = V_{DD}$ and I_H is the drain current at $V_{GS} = V_{DD}$ and $V_{DS} = V_{DD}/2$. While this will not be exactly applicable to modeling the dynamic switching current of ring oscillator, it implies that IDDA is almost solely determined by the I-V relationship. Therefore, one should expect that measured IDDA is captured by SPICE simulations when the DC model is plugged in directly, regardless of the assumed wire load which may affect the predicted delay. When the supply voltage is increased, the bias voltages corresponding to I_{eff} is also increased, which allows for probing of the fast I-V response at various bias combinations.

The gate stage delay (τ) is characterized as the RO oscillation period divided by the number of stages (101). Note that each oscillation period consists of 202 switching events because each internal node encounter one rising and one falling edge signal. Therefore, the calculated τ equals to the summation of the characteristic delays of inverter node charging through PFET and the discharging through NFET. The delay should be much higher than the NC response time for the validation to be successful at the corresponding operation condition.

Effective Resistance and Capacitance

During the active operation, only one stage is switching at a time, while the rest still draw static leakage current from the supply. Therefore, the effective switching current is defined as $IDDA - IDDQ$, and it can be used to characterize an effective resistance (R_{eff}) for the circuit switching

$$R_{eff} = V_{DD}/(IDDA - IDDQ) \tag{2.2}$$

The effective load capacitance is then defined as

$$C_{eff} = \frac{\tau}{R_{eff}} = \frac{\tau(IDDA - IDDQ)}{V_{DD}} \tag{2.3}$$

Modeling of τ requires knowledge of the unknown wire load, and hence direct comparisons between the simulated and measured delay is inappropriate for the validation purpose. However, the derived C_{eff} is an appropriate quantity for validating the fast response, as it carries valuable information of transistor gate capacitance in its slope with respect to V_{DD}. To the first order, C_{eff} can be estimated as the average input capacitance of an FO3 stage over the bias range (0V to V_{DD}) at the gates. As the supply voltage increases, the ratio of V_{th}/V_{DD} reduces for both the NFET and PFET, and thus the relative bias regime under which the device is in strong inversion is extended, resulting in a larger overall C_{eff}. Therefore, the increase of C_{eff} as a function of V_{DD} reflects the magnitude change of the MOSFET gate capacitance from sub-threshold regime to strong inversion. Besides, the slope also accounts for second order effects such as DIBL which reduces the V_{th} at large V_{DD}. Therefore,

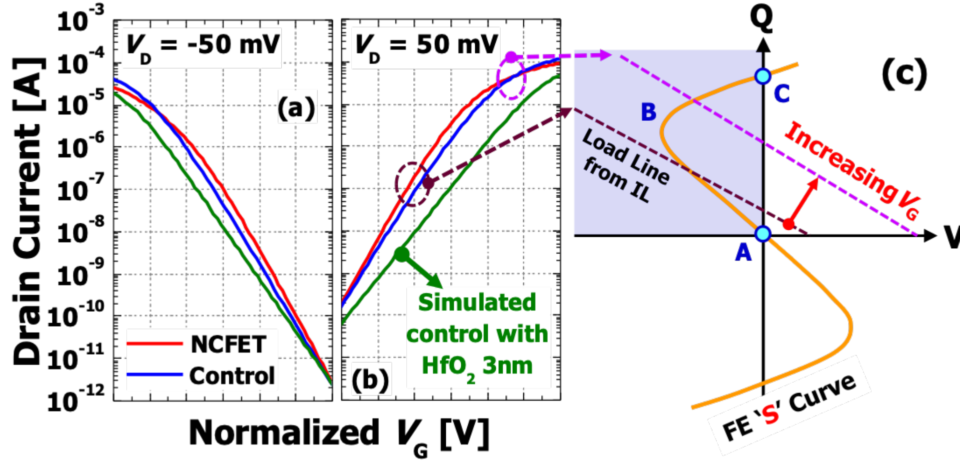


Figure 2.2: (a) V_{th}-matched I_D-V_G characteristic of Control MOSFET and NCFET. For comparison, simulated characteristics of devices with 3 nm DE is included. (b) A schematic 'S' curve of the ferroelectric showing the region of operation.

it is non-trivial to match the extracted slope to the simulation unless the capacitive response of the gate at the RO frequency is consistent with the compact model. A good agreement between the measurements and the quasistatic model would also imply insignificant gate resistance, which may otherwise indicate a large dissipative term (ρ) of the LK equation.

2.3 Experiments and Models

State-of-the-art replacement gate 14nm FinFET devices with doped HfO2 ferroelectric layer of 3 nm thicknesses were fabricated in a GlobalFoundries process. 101 stage FO3 ring oscillators were used to study RO performance. The highest voltage to be applied (lowest delay measured) was chosen so that it does not damage the devices.

NCFET Characteristics

Representative DC characteristics of an NC-NFET and NC-PFET with that of a control with much thinner dielectric layers are shown in Fig. 2.2(a),(b). The steepening of the subthreshold swing is clearly evident for both cases despite a thicker gate stack. Note that, for comparison, simulation of a NFET and PFET with 3 nm gate dielectric (DE) are also included, where the thick DE thickness would lead to significantly degraded swing and on current. Fig. 2.2(c) shows the regions of operation in terms of the ferroelectric 'S' curve. At small V_{GS}, the load line is in the NC region, leading to steeper swing. At large V_{GS}, a large capacitance can still be measured, while the NCFET ON current is lower than the Control. Authors of [28] term such devices as 'PCFET', postulating that the NC-NFET operation

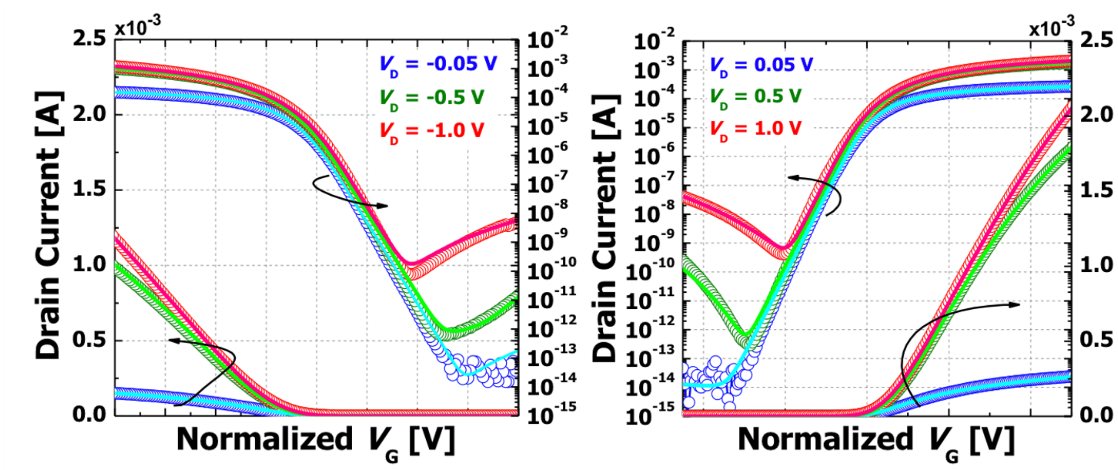


Figure 2.3: Measured and modeled I_D-V_G for (left) an n-type NCFET and for (right) a p-type NCFET. Both semi-log scale and linear scale current are plotted. Dotted and solid lines represent measurements and modeling, respectively.

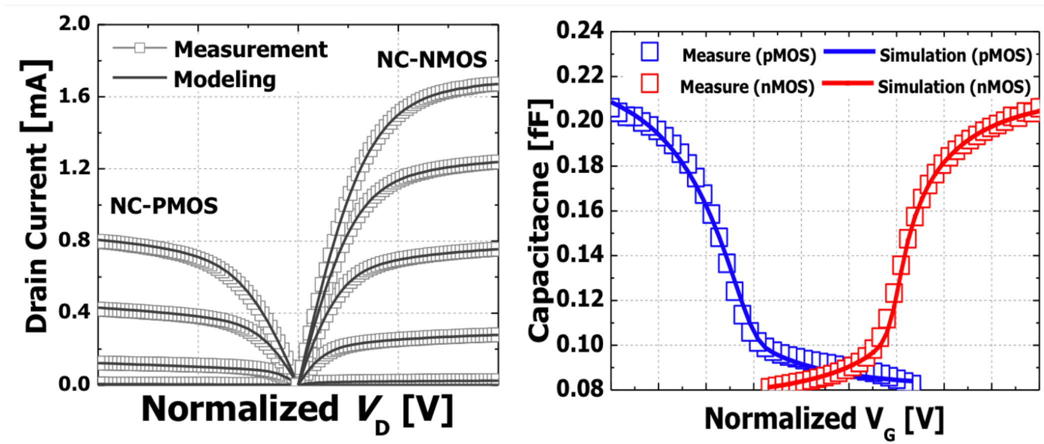


Figure 2.4: Measured and modeled (left) I_D-V_D at various V_{GS} and (right) C-V at $V_{DS} = 0V$ for both the NC-NFET and NC-PFET.

traverses in the “S” curve following ABC-CBA in Fig. 2.2(c). Note that the lower current for the NCFET may also be a result of extra source and drain underlap due to the larger side wall thickness for the FE layer. Nonetheless, the device shows hysteresis-free behaviors with ferroelectric gate stack, indicative of the negative capacitance response.

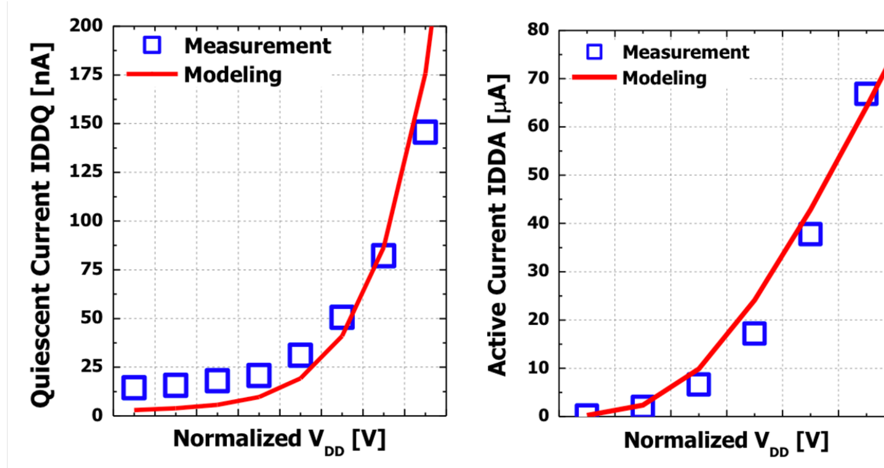


Figure 2.5: Simulated and measured (a) IDDQ and (b) IDDA on the 101 stage FO3 RO. The simulation was performed with quasistatic models calibrated to the DC measurements.

Quasistatic Device and Circuit Model

The I_D-I_G, I_D-V_D, and C-V characteristics of both the NC-NFET and NC-PFET are measured from large FinFET arrays and calibrated to BSIM-CMG compact models (Fig. 2.3 and 2.4.) Excellent agreement between the model and the measurement is obtained. The calibrated models are then used to simulate the 101-stage FO3 RO circuit through HSPICE without any NQS modifications to predict IDDA, IDDQ, and C_{eff}.

RO Characterizations

Fig. 2.5 shows simulated and measured IDDQ and IDDA. The IDDQ at low V_{DD} is underestimated by the simulation probably because of slight V_{th} variations of the RO devices and the sensitive nature of OFF current to the threshold voltage. However, the increasing trend of IDDQ with respect to V_{DD} is captured by the the model, indicating that the DIBL and the GIDL effects, and hence the device electrostatics, are consistent between the model and the RO devices. The agreement of IDDA between simulation and measured data is a critical result as it shows that IV and CV characteristics remain the same at these high frequencies as DC. To resolve possible controversy that the matched results may not correspond to the NC response but instead the “PC” region at high biases as discussed previously, we take a close look at the voltage range where the maximum value of V_{DD} is low enough so that the PFET is still in the NC region (Fig. 2.6). This ensures that as the gate voltage goes up and down the RO will always be in the NC region for at least one type of devices. The IDDA simulated at such V_{DD} closely matches the measured values, which confirms the efficacy of the DC model in this low voltage region.

Next, we look into the calibration for the RO stage delay and derived C_{eff} (Fig. 2.7).

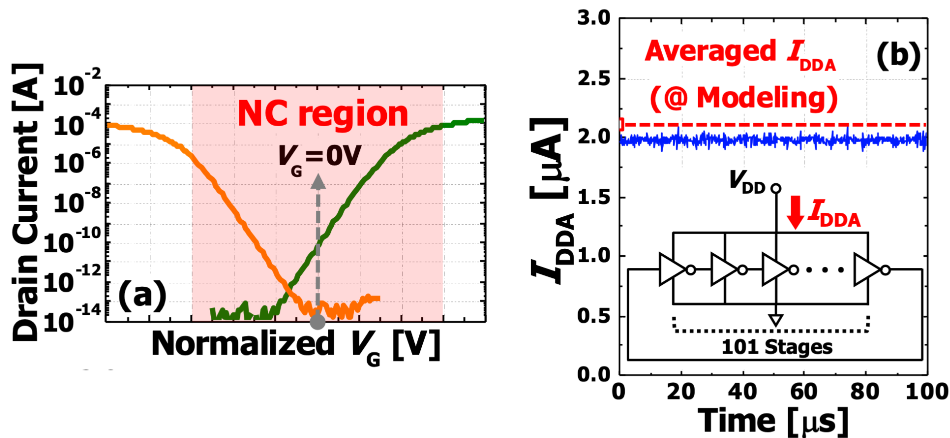


Figure 2.6: (a) Operation range corresponding to a V_{DD} where the PFET remains in NC regime. (b) Simulated and measured IDDA at the V_{DD} indicated in (a).

When the RO model is simulated without wire capacitance (C_{wire}), the model would predict C_{eff} that is consistently 0.4fF lower than the experimental results over a wide range of V_{DD}. The significant observation from this comparison is that the slope of the C_{eff}-V_{DD} relationship from the experiment agrees with the model, showing that the variable capacitive response of the gate is captured by the model at high speed. Note that the C_{eff} increasing linearly with V_{DD} is also a signature of the fact that the devices can respond to input signal changes at the oscillation frequency. A good fit to the C_{eff} can therefore be achieved by simply adding a 0.4fF C_{wire} to the fanout of each RC stage, which indeed shifts the C_{eff}-V_{DD} up by C_{wire} without affecting the slope. It worth emphasizing that the C_{wire} added to the model affects the IDDA negligibly. Having achieved good matches for IDDA, IDDQ, and C_{eff}, the RO stage delay is therefore successfully reproduced by the model.

RO Speed and Projections

The device model is calibrated to cover the bias ranges of RO simulations with V_{DD} under the shaded area in Fig. 2.8. As a result, the matching between the modeling and the experiments indicates that there is no slowdown of NC effect at a gate stage delay (including both NMOS and PMOS switching) slightly lower than 20 ps. Beyond the fitted regime, the measured RO delay keeps decreasing and is lower than the simulation results. While this is outside of the range of direct validation, we point out that this trend is consistent with conventional RO operations, and we didn't observe discernible extra delay that can be ascribed to the NC response for gate stage delay down to 8.5 ps.

We also note that the RO is made with only n-type work function gate metal. As a result, the V_{th} for the NC-PFET is significantly larger than the NC-NFET. To see the effect of this mismatch, we have simulated the ROs with the same device parameters but by artificially

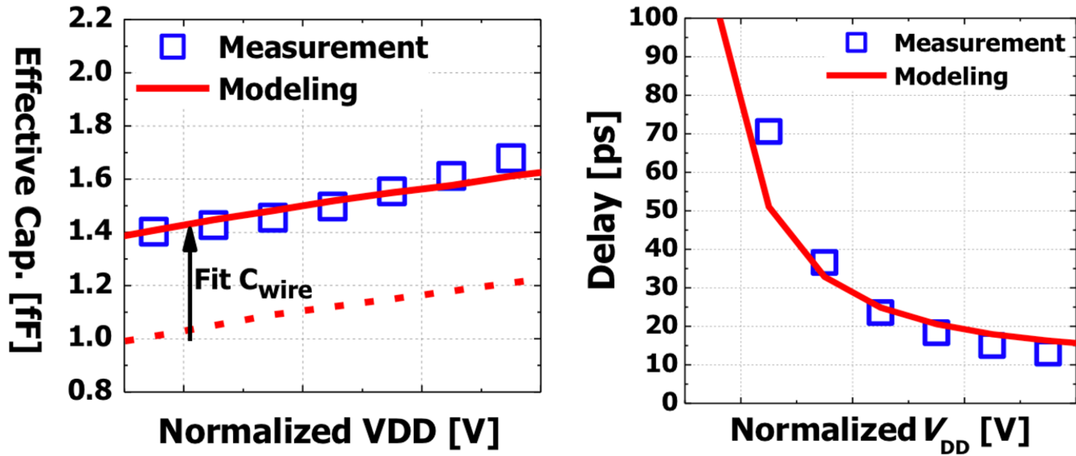


Figure 2.7: The simulated and measured (left) C_{eff} and (right) gate stage delay τ for the FO3 NCFET RO. (Left) The dashed line is simulated without wire load, and the solid line is simulated with 0.4fF wire load. (Right) The delay is reproduced by the model with 0.4fF wire load.

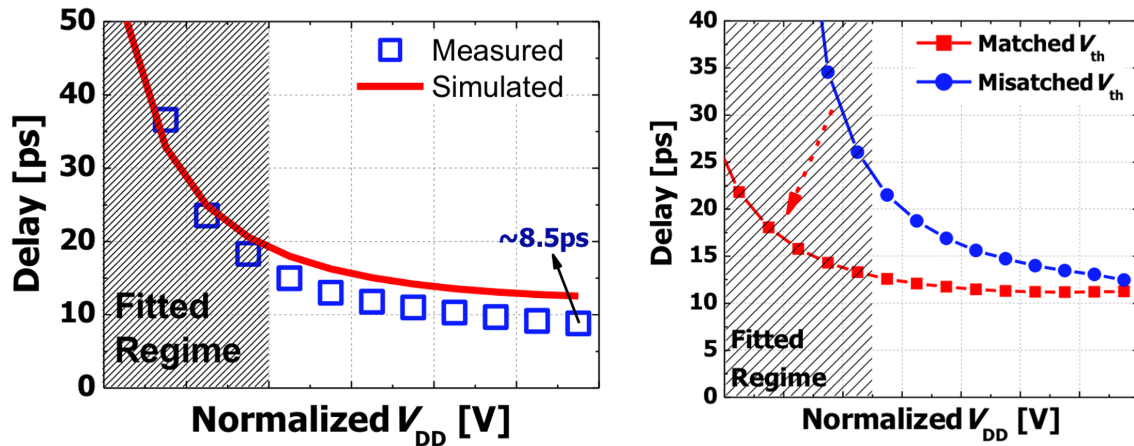


Figure 2.8: (Left) Simulated and measured gate stage delay. The shaded area stands for the V_{DD} range under which the device model is calibrated to DC measurements. (Right) Simulated gate stage delay and the projection with matched V_{th}. The shaded area stands for the trustworthy regime for matched-V_{th} simulations.

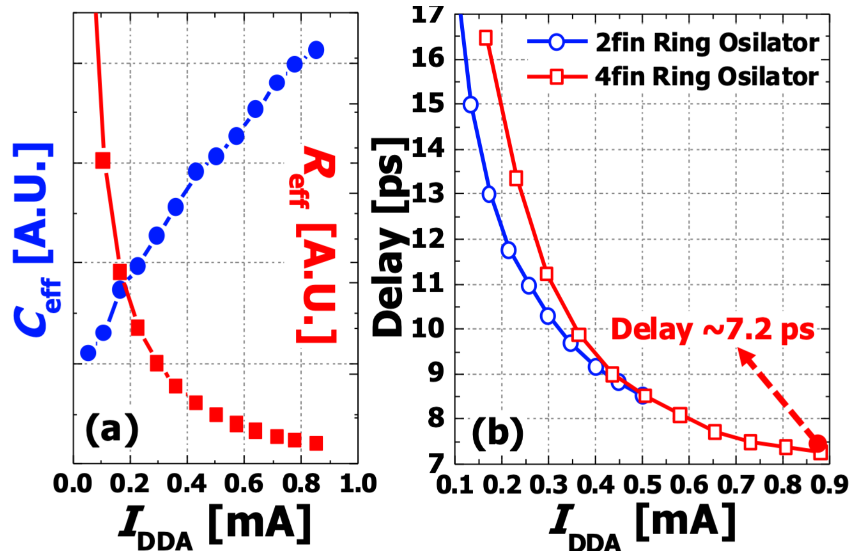


Figure 2.9: (a) Measured C_{eff} and R_{eff} from 4-fin NCFET ring oscillator. Behavior of C_{eff} and R_{eff} follow expected trends. (b) Delay vs. IDDA curves for 2 fin and 4 fin NC-FinFET ring oscillators. Stage delay as small as 7.2 ps was measured.

increasing the PFET gate workfunction to match the V_{th}. This significantly improves delay in the low V_{DD} region (Fig. 2.8). At the very high speeds, the effect is smaller, indicating that it is the series resistance and parasitic capacitance that limits the highest speed (8.5 ps) rather than the intrinsic device. Finally, we measured ROs made of 4 fin devices that provide larger ON current. The ROs behave similarly to 2 fin ROs and a gate delay as small as 7.2 ps can be measured (Fig. 2.9).

2.4 Conclusions

We have demonstrated that the behavior of the RO based on 14nm NCFETs can be predicted from DC characteristic of the individual devices, just like conventional CMOS technology, thereby demonstrating that there is no extra slowdown attributable to NC effect even at the very high frequencies used in this study. Stage delay down up to 7.2 ps has been measured for FO3 inverters and limited by resistance and capacitances of the circuit as usual. Therefore, well optimized NCFETs should operate at very high frequencies at reduced power dissipation.

Chapter 3

Modeling of the Anomalously Beneficial Gate-Length Scaling Trend

3.1 Introduction

In this chapter, we show that the nonlinearity of NC effect, stemming from the S shape of the polarization-electric field relationship, gives rise to an unconventional trend of gate length (L_g) scaling that is not expected from conventional MOSFET theory. This effect is a distinctive behavior of negative capacitance transistors. The increase of ferroelectric polarization due to the inner fringing field [44, 69] contributes to the L_g-dependent potential amplification effect, which translates to gate stack capacitance improvement in short-channel NCFETs. This trend is exactly the opposite of what is seen in conventional MOSFETs, where decreasing L_g leads to worsening of the gate capacitance. By comparing reference devices and NCFETs across many gate lengths, we show this effect experimentally and confirm it with TCAD numerical simulations. [1]

3.2 Sub-threshold Scaling Trends

The device characteristics in the sub-threshold regime is well described by the equivalent capacitance model [70]. At room temperature,

$$SS \equiv \left(\frac{\partial \log_{10} I_D}{\partial V_G} \right)^{-1} = m \times 60(mV/dec) \tag{3.1}$$

where m is the body factor. For an ultra-thin-body (UTB) short channel device,

$$m = \frac{\partial V_G}{\partial \phi_s} = 1 + \frac{C_s + C_d}{C_g} \tag{3.2}$$

[1] The content of this chapter is based on [59].

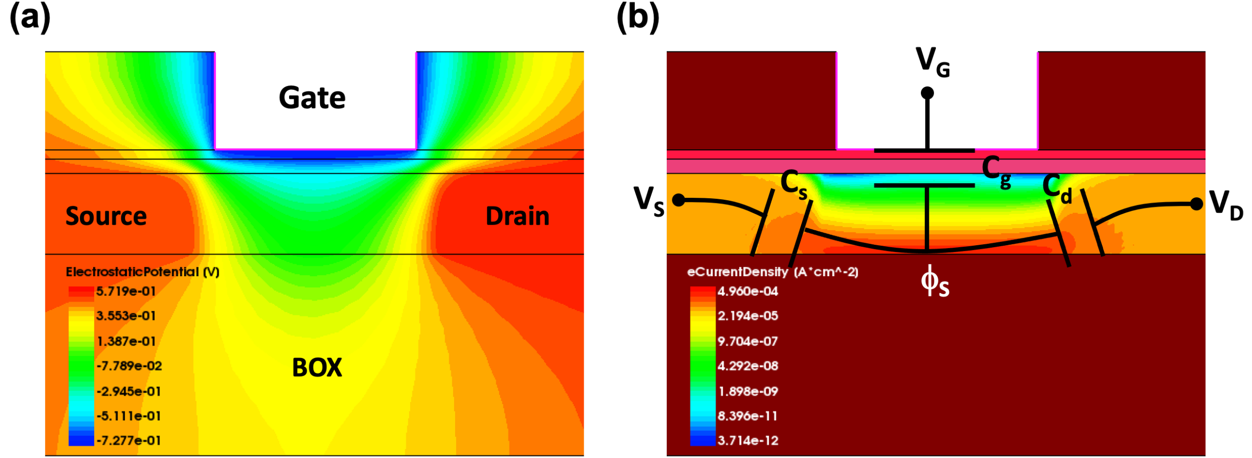


Figure 3.1: (a) Electrostatic potential profile and (b) current density profile of a short-channel n-type SOI MOSFET biased in sub-threshold regime ($V_{GS} = 0V$ and $V_{DS} = 0.05V$.) Because of the built-in potential between the gate and the source and drain, the electrostatic coupling causes the potential to be lower at the channel surface than at the bottom. As a result, the leakage is dominated by the bottom path, which implies the TOB best describes SS and DIBL is located at the bottom of the Si layer.

where ϕ_S is the electrostatic potential and C_s, C_d, and C_g are differential capacitance from the top of the energy barrier (TOB) to the source, drain, and gate, respectively. Similarly,

$$DIBL \equiv -\frac{\partial V_G}{\partial V_D}\Big|_{I_D} = \frac{C_d}{C_g} \tag{3.3}$$

To the first order, changing the gate stack does not affect C_s and C_d. Therefore, for two MOSFET devices with different gate stacks but the same geometry, $m - 1$ and DIBL ratios are dominated by the ratio of C_g.

$$\frac{DIBL_1}{DIBL_2} \approx \frac{m_1 - 1}{m_2 - 1} \approx \frac{C_{g2}}{C_{g1}} \tag{3.4}$$

Conventional Trend for EOT Scaling

For electrically well-tempered transistors, the sub-threshold current is dominated by the leakage path near the channel surface. In this case, the TOB is located at the near channel surface, and therefore C_g is close to the constant capacitance of the gate stack ($C_{stack} = \epsilon_{SiO2}/EOT_{stack}$.) For un-doped UTB transistors, little band bending exists in the channel in the sub-threshold regime, which implies that the current paths at different channel depths contribute similarly to the leakage. For each path across the depth, the current can be

describe by a top-of-the-barrier model, and the overall m and $DLBL$ is an average for all the paths.

However, in a short-channel device, a considerable band bending can exist in the subthreshold regime (Fig. 3.1(a)) because of the electrostatic coupling to the source and the drain. The electrostatic potential at the surface is lower than that at the bottom. Equivalently, the conduction band energy at the surface is higher than that at the bottom, leading to relatively increased leakage at the bottom. As the short-channel effect (SCE) is aggravated, the sub-surface leakage path at the bottom starts to dominate the leakage (Fig. 3.1(b).) Therefore, the weight of the bottom leakage path towards the overall $DIBL$ and m is boosted significantly. As a result, the TOB is moved from the interface into the silicon body, and an extra segment of silicon body from of the SOI thickness directly degrades the C_g (Fig. 3.1(b).) This not only decreases gate capacitance but also reduces the C_{g2}/C_{g1} ratio for device 1 and device 2 of the same geometry with $EOT_2 < EOT_1$. Therefore, for conventional MOSFETs, the improvement ratios of $m-1$ and DIBL due to a thinner EOT gate stack diminish as L_g scales down.

Anomalous Experimental Results for NCFET

By contrast to the conventional electrostatics, we observe the opposite trend in the experiments done on NCFETs and high-κ control transistors. NCFETs and high-κ (to be referred to as the control) transistors were fabricated using FDSOI wafers. Fabrication steps and processes (including thermal processing) are identical for both transistors except the NCFET wafer has Zr doping in what would otherwise be a pure HfO_2 layer. Detailed process flows and characterization results are specified in [34]. Figure 3.2 summarizes the measurement results at 300K of the control devices and NC devices. The NCFETs show a larger improvement of SCE, and the improvement ratios of $m-1$ extracted from measured SS and DIBL increase with decreasing L_g. As described above, this trend of improvement cannot be simply attributed to a thinner EOT. The C_{stack} of NCFETs appears to increase for short gate-length devices where a large C_{stack} is needed the most.

3.3 Modeling of the NCFET Scaling Trends

Relationship Between Polarization and Gate Capacitance

For a stabilized FE material with first-order phase transition, $E_{FE} = 2\alpha P_{FE} + 4\beta P_{FE}^3 + 6\gamma P_{FE}^5$ where $\alpha < 0$ and $\gamma > 0$. Accordingly, the capacitance of FE layer with thickness t_{FE} and dielectric constant ϵ_r is

$$C_{FE} = \left(\frac{1}{2\alpha + 12\beta P_{FE}^2 + 30\gamma P_{FE}^4} + \epsilon_r \epsilon_0 \right) \frac{1}{t_{FE}} \tag{3.5}$$

which can be negative for small P_{FE}. Note that neglecting high order (β and γ) terms would imply constant negative capacitance, which is not assumed in this work. The gate stack

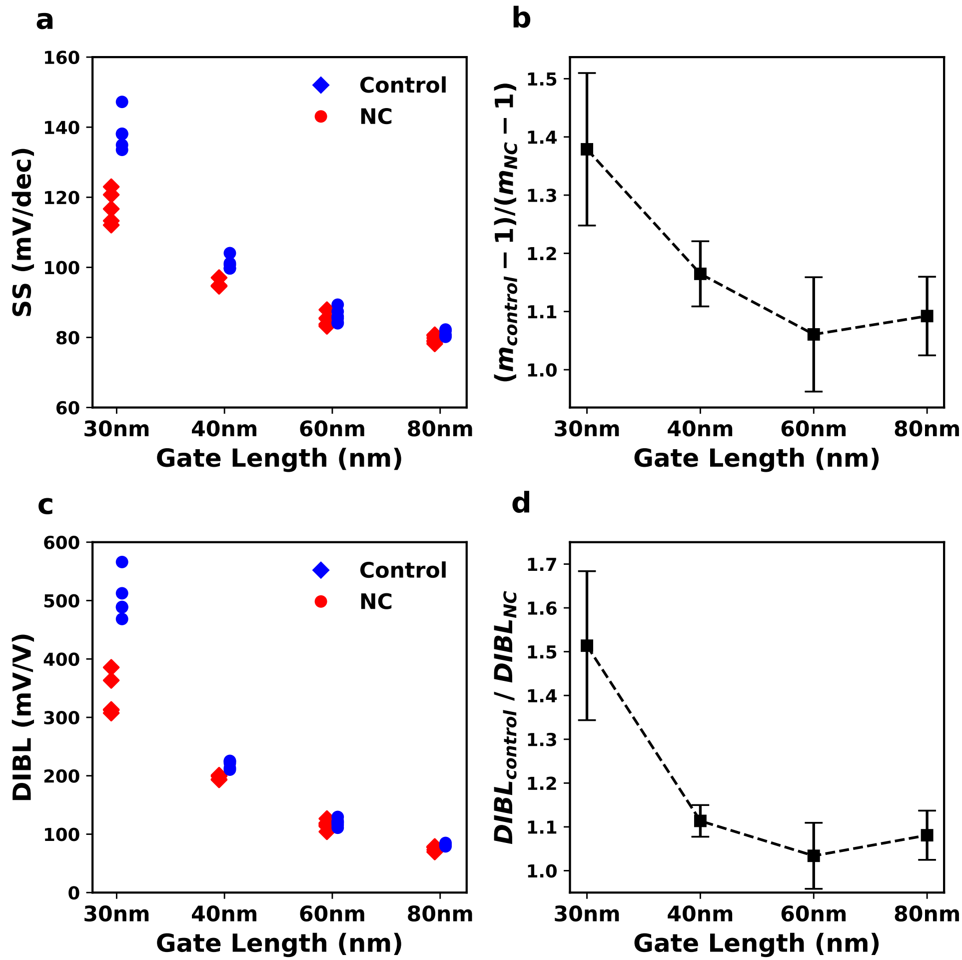


Figure 3.2: Extracted (a) SS scatter plot, (b) control-to-NC $m - 1$ ratio with error bar, (c) DIBL scatter plot, and (d) control-to-NC DIBL ratio with error bar, measured for different L_g devices. Each data point in (a) and (c) corresponds to one measured device, where a small horizontal shift is used to separate control and NC results for better visualization. The bars in (b) and (d) indicate the sample standard deviation of each ratio due to device-to-device variations.

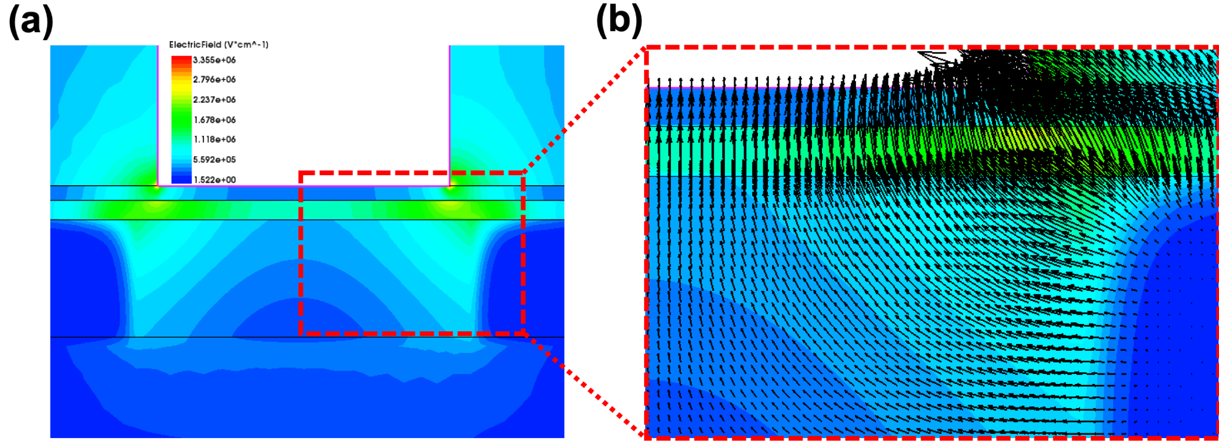


Figure 3.3: Electric field profile of a short channel SOI MOSFET in sub-threshold regime. The light color in the silicon channel near the source and the drain edge is due to the inner fringing field coupling, and it does not fully decay at the middle of the channel. (b) is a zoomed-in view of (a) from mid-channel region to the drain edge with arrows indicating the field direction. The electric field on the gate stack is nearly vertical for the mid-channel region.

capacitance, which forms the major component of C_g, can be expressed as

$$C_{stack} = \left(\frac{1}{C_{FE}} + \frac{1}{C_{IL}} \right)^{-1} \tag{3.6}$$

where C_{IL} is a positive capacitance due to the interfacial layer. A less negative C_{FE} gives rise to a larger C_{stack} and C_g.

Relationship Between Inner Fringing Field and Polarization

The SCE is fundamentally associated with the increase of electrostatic coupling between the channel and the source and drain, which is characterized as C_s and C_d in the TOB model. In terms of electric field, the C_d and C_s is due to the inner-fringing-field coupling that is the strongest near the source and the drain edges and gradually decays in the channel (Fig. 3.3.) For devices with shrunk L_g, the inner fringing field may not fully decay in the middle of the channel, which has two consequences for UTB devices.

- Increased C_s and C_d. Unless C_g is negative, this implies worsened SCE.
- Finite external vertical electric field at the gate stack below which the TOB resides. The shorter the L_g, the less decayed the fringing field near the TOB, and thus the larger the external field on the gate stack.

Therefore, for UTB devices in the sub-threshold regime, reducing L_g increases the vertical field on the gate stack near the TOB, which further polarizes the P_{FE} that affects C_g [69]. Now, we identify two possible explanations for the anomalous scaling trends based on the S-curve model of negative capacitance.

Model 1: First-Order Ferroelectricity

For ferroelectric material with first-order phase transition, the fourth-order energy coefficient in the Landau model (β) is negative. As $|P_{FE}|$ is small and β is negative, increased $|P_{FE}|$ due to the inner fringing field would increase the abvsolute value of the already negative-denominator of the first term in the RHS in eq. 3.5. Therefore, reduced L_g may boost the FE polarizability and results in less negative C_{FE} and thus larger C_{stack}.

Model 2: Negative Fixed Interface Charges

It has not been settled in the literature whether the FE phase transition of thin-film Zr-HfO_2 is first-order or second-order. Therefore, we propose another model that relaxes the assumption of negative β. Here we assume negative fixed charges at the SiO_2/FE interface, which induces a built-in electric field in the FE that polarizes large P_{FE} in the sub-threshold regime. In such systems, the inner fringing field would counterbalance the built-in oxide field and reduce $|P_{FE}|$. If the $|P_{FE}|$ in long channel devices is large so that the positive highest-order (γ) term in eq. 3.5 dominates the P-E relations, a reduced $|P_{FE}|$ will result in a strengthened NC effect. For both types of NCFET systems, reducing L_g results in enhanced C_{stack} and potentially increases the $m - 1$ and DIBL improvement ratios.

3.4 Simulation Results

The NCFETs with Control experiments are self-consistently modeled according to the methodology outlined in Fig.1.5. Sentaurus TCAD [25] is used to numerically investigate the scaling trends. To calibrate with experiment, FDSOI structures with 12nm silicon thickness and thick buried SiO_2 are simulated, with L_g varying from 30nm to 80nm and the source and drain doping profile fixed. Control devices (EOT=2.31nm), "higher-κ" conventional devices (EOT=1.73nm), and NCFETs, all having 3.5nm physical thickness, are simulated. The NCFET gate stack is composed of a 2nm SiO_2 interfacial layer and a 1.5nm ferroelectric layer. For NCFETs, the Landau-Khalatnikov (L-K) equation with a negligible polarization gradient energy term for ferroelectric polarization is solved self-consistently with Poisson and transport equations. The resultant FE polarization profile changes gradually along the channel in accordance with the external electric field. If a larger polarization gradient energy term were adopted, the inner fringing field near the source and drain would have a larger effect on the channel electrostatics [41], resulting in an even larger SCE improvement and stronger anomalous scaling trend. For all the simulated I_D-V_G relations, SS is averaged over

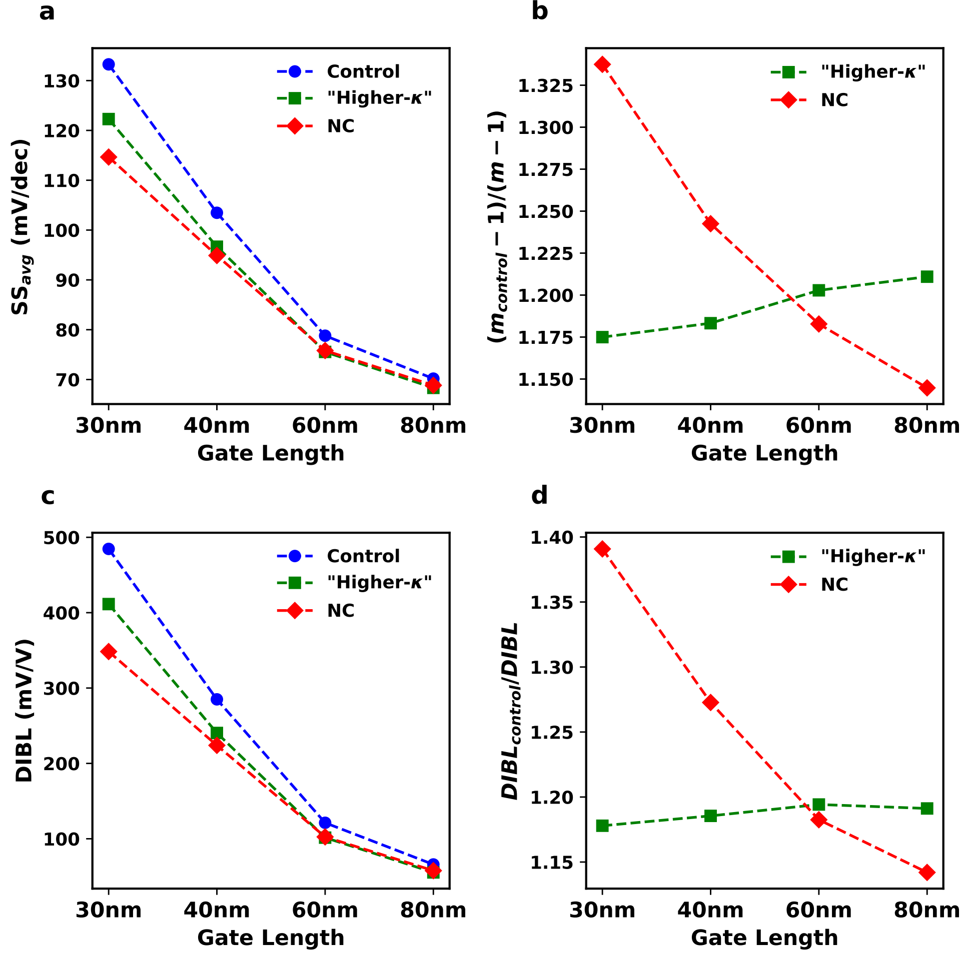


Figure 3.4: Simulated (a) SS, (b) $m-1$ ratios of control devices to improved devices, (c) DIBL, and (d) DIBL ratios of control devices to improved devices, as a function of L_g. The control devices consist of 2.31nm-EOT gate stacks. Improved devices include "higher-κ" conventional devices (EOT=1.73nm) and NC devices ($\alpha = -2.4 \times 10^{11} cm/F$, $\beta = 0 cm^5/(FC^2)$, $\gamma = 3.2 \times 10^{33} cm^9/(FC^4)$, $\epsilon_r = 19$, and SiO_2/FE interface charge density of $-0.64\mu C/cm^2$).

4 decades of drain current in the deep sub-threshold regime, and DIBL is extracted as the constant-current (1nA/μm) gate voltage change between V_{DS}=50mV and V_{DS}=1V.

Reproduction of Experimental Trends

Both "higher-κ" and NC devices present better SS and DIBL than the control devices. However, the higher-κ simulations have more severe SCE than the NC simulations for $L_g \leq$

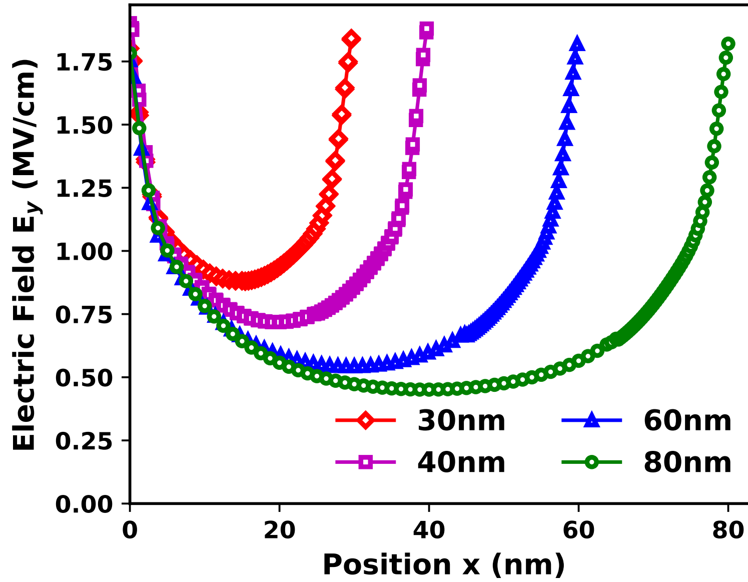


Figure 3.5: Simulated out-of-plane electric field (E_y) in the interfacial layer of NCFETs with different L_g as a function of position along the channel (x) from source edge ($x = 0$) to drain edge ($x = L_g$) at $V_{GS} = 200mV$ and $V_{DS} = 50mV$ (in sub-threshold regime.)

$40nm$ despite having better gate control for $L_g \geq 60nm$ (fig. 3.4 (a), (c)). The improvement ratios relative to the control devices (fig. 3.4 (b), (d)) show opposite trends, where NCFETs demonstrate increased C_g ratio as opposed to conventional MOSFET scaling theory. Both the "higher-κ" and NC scaling trends confirms the above discussions.

Discussions

Figure 3.5 confirms that the inner fringing field near the top of the barrier increases as the L_g of the NCFET shrinks. Figure 3.6 shows the simulated relations between gate charge density and ferroelectric voltage at the mid-point between source and drain, which is known to be close to the top of the energy barrier. If fixed charges of $-0.64\mu C/cm^2$ exist at the SiO_2/FE interface, Q_G is approximately $0.64\mu C/cm^2$ in sub-threshold for an 80nm device (Fig. 3.6 (a)). Scaling down of L_g reduces Q_G as well as P_{FE} in sub-threshold, transitions the FE from the positive-capacitance to negative-capacitance regime, and increases C_{stack} (Fig. 3.6 (b)). When no fixed charges exist at the interface, Q_G and hence P_{FE} are close to zero in sub-threshold when $L_g = 80nm$ (Fig. 3.6 (c)). When L_g is reduced, the increased inner fringing field drives a larger $|P_{FE}|$. This leads to a decrease of $|C_{FE}|$ and increase of C_{stack} when β is negative (Fig. 3.6 (d)). Both mechanisms manifest as enhanced potential amplification and gate control improvement as shown in Fig. 3.4.

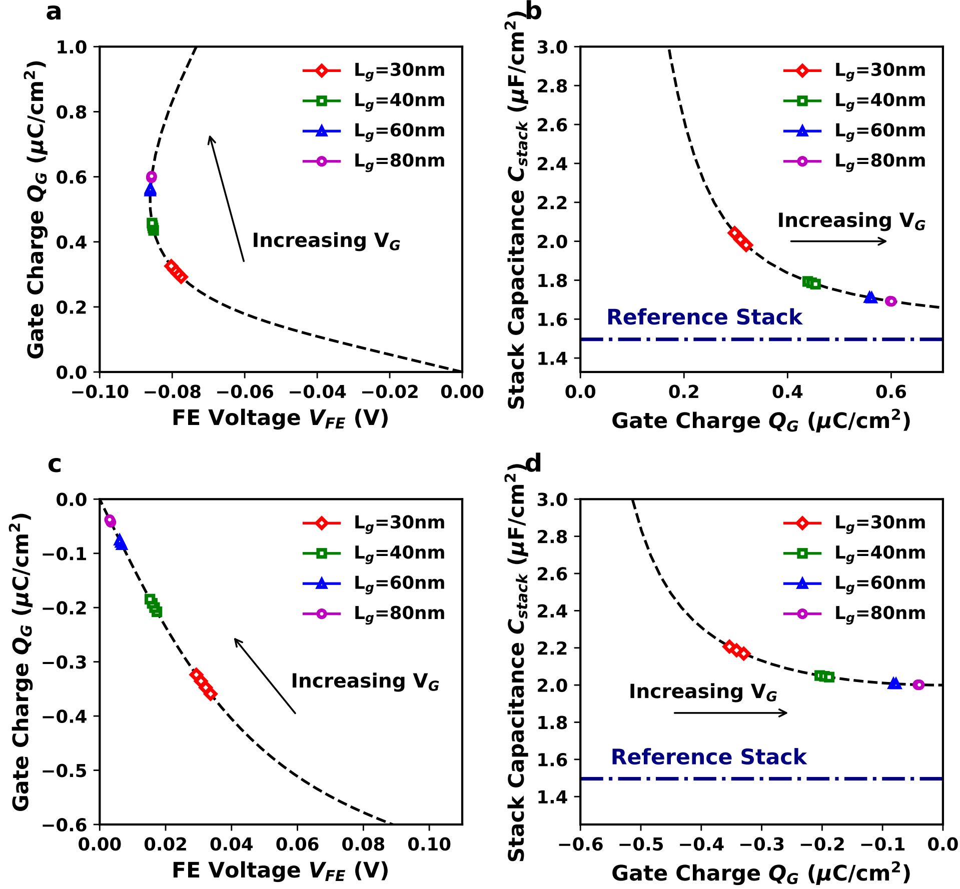


Figure 3.6: Simulated (a), (c) $Q-V$ relations of the ferroelectric layer and (b), (d) $C-Q$ relations of the gate stack at the mid-channel of NCFETs with different L_g. For each L_g, the gate voltage is ramped at which I_D is from 10pA/μm to 100pA/μm at V_{DS}=50mV. For (a) and (b), the ferroelectric parameters are $\alpha = -2.4 \times 10^{11} cm/F$, $\beta = 0 cm^5/(FC^2)$, $\gamma = 3.2 \times 10^{33} cm^9/(FC^4)$, and $\epsilon_r = 19$. The devices in (a) and (b) are simulated with $-0.64\mu C/cm^2$ fixed charges at the SiO_2/FE interface. For (c) and (d), the ferroelectric parameters are $\alpha = -1.4 \times 10^{11} cm/F$, $\beta = -1.2 \times 10^{22} cm^5/(FC^2)$, $\gamma = 8.0 \times 10^{32} cm^9/(FC^4)$, and $\epsilon_r = 19$. No interface charges are assumed in (c) and (d).

We also investigated an imaginary system where the FE material has nonlinear but positive capacitance, corresponding to positive α and β in the L-K equation. The nonlinear positive capacitance layer would lead to neither a significant improvement of C_g nor observable scaling trends (Fig. 3.7). The simulations confirm that the anomalous trend is a signature of the nonlinear NC effect.

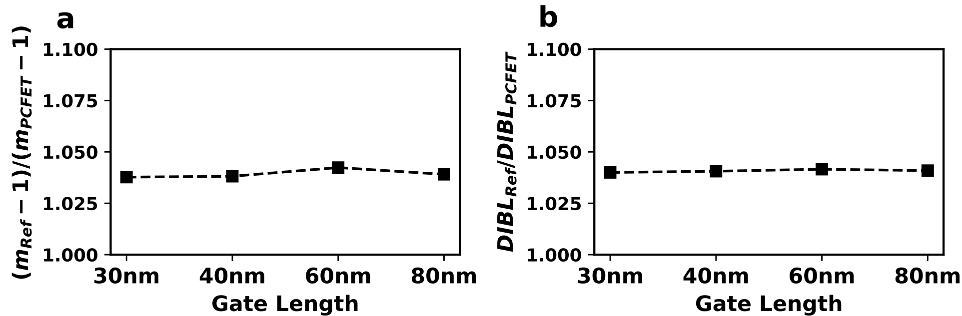


Figure 3.7: Simulated (a) $m - 1$ ratios (b) DIBL ratios between control MOSFETs and ferroelectric FET with nonlinear positive-capacitance oxide. For the imaginary positive-capacitance ferroelectric material, $\alpha = 4 \times 10^{11} cm/F$, $\beta = 10^{22} cm^5/(FC^2)$, and $\gamma = 10^{31} cm^9/(FC^4)$.

3.5 Conclusion

Improved gate stack capacitance with decreasing channel length, opposite to the trend seen in conventional MOSFETS, was observed in both simulations and experiments of NCFETs. The results can be explained by the nonlinearity of negative capacitance and the increase of the inner fringing field due to gate length reduction. This unique gate length dependency of gate control provides a new way to distinguish NCFETs from conventional MOSFETs.

Chapter 4

Modeling of the Near-Threshold Swing Reduction Effect

4.1 Introduction

The objective of reducing supply voltage can be achieved by improving the swing near the threshold, and this can be done without violating the condition necessary for zero hysteresis [50]. NCFET so designed, even if does not show sub-thermal SS, can show other subthreshold behavior that are not expected classically. Understanding such behavior quantitatively requires accounting for the non-linearity of the S-shaped Polarization-Electric Field (P-E) relation [71, 72, 73, 74, 75] together with non-uniformity of polarization along the channel that arises due to two-dimensional electric field. Specifically, the degree of suppression of the polarization and non-uniformity (domain formation) that play a critical role in the stabilization and strength of the negative capacitance [54, 18, 13, 76, 24] will determine the specific behavior observed in the current-voltage characteristic.

In this chapter, we present combined experiment-model study on Silicon-on-Insulator (SOI) short-channel NCFET devices that demonstrates the non-classical behavior and explains the origin of it. The ferroelectric gate material is modeled using the Landau and Devonshire theory [10] , which is solved self-consistently with the electrostatics and carrier transport through the TCAD simulator to quantitatively reproduce the observed differences between the Control experiments and the NCFET for multiple gate lengths (L_G) and drain biases (V_{DS}.) This chapter is organized as follows. In 4.2 we discuss the experimental characterization of the Control and NCFET devices and identify the non-linear characteristics of the NCFET that are not expected conventionally. In section 4.3 we present the calibration procedure of the model to the experiments with the aid of TCAD. The physical insights behind the gate stack non-linearity in terms of both the L_G and V_D scaling as well as other implications from the model are discussed in section 4.4, followed by a brief conclusion in section 4.5. [1]

[1] The content of this chapter is based on [60].

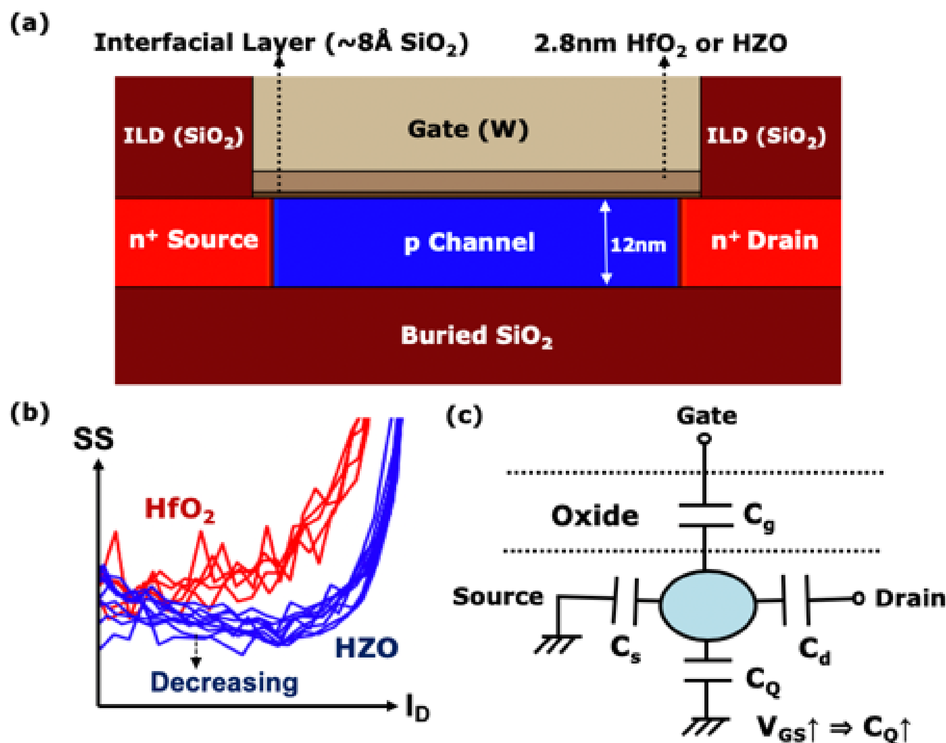


Figure 4.1: (a) Schematic cross-section of the SOI n-MOSFETs. Both HfO_2 (Control) and HZO (NC) devices have the same geometry. (b) SS of HfO_2 and HZO devices demonstrating conventional and unconventional trend with respect to I_D, respectively. (c) Equivalent circuit diagram of the MOSFET in sub-threshold regime.

4.2 Device Characterizations

Experiments

Fig. 4.1(a) shows the schematic of the device being used in this study. NCFETs are fabricated using laminated, atomic layer deposited layers of HfO_2 and ZrO_2 (HZO.) In comparison, Control devices are made of HfO_2 only. The thickness of the layer in both cases is 30 ALD cycles. Details regarding the device fabrication and characterizations are discussed in [35]. In both cases, a SC1 induced 'chemox' interfacial layer (IL) is used (~8A.) The IL synthesis is common between both groups of transistors, and all other processing and annealing conditions are also the same. For both groups, W is used as the gate metal. The gate stack consists of no metal interlayer between the HZO and the IL. Substantially improved I-V characteristics can be observed for NCFETs compared to the Control, systematically over multiple channel lengths (Fig. 4.1(b) and 4.3, and 4.4.) Particularly interesting is the sub-threshold behavior where the NCFETs show improving sub-threshold swing with increasing

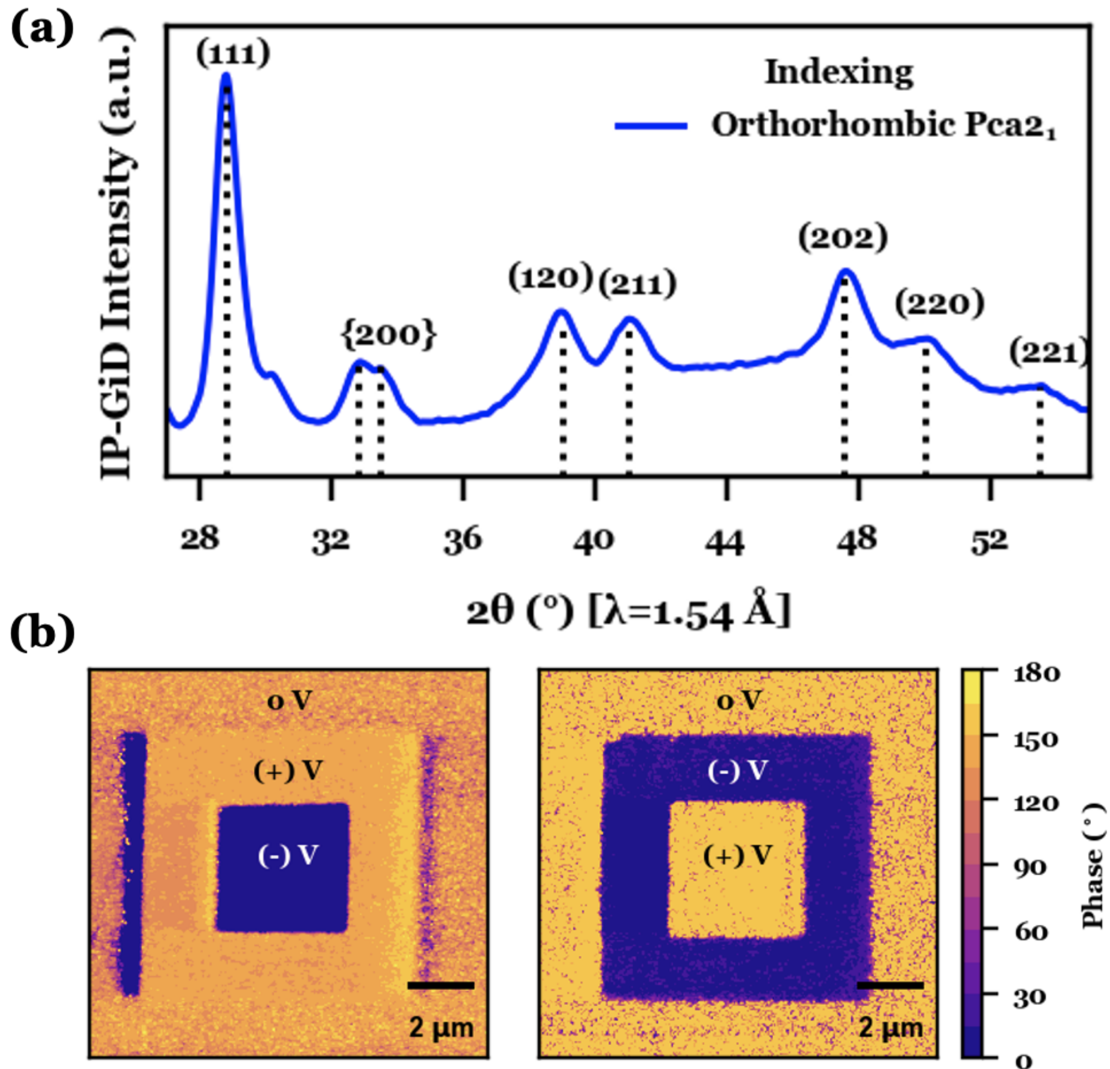


Figure 4.2: Demonstration of inversion symmetry breaking and switchable electric polarization in 3 nm ZHO film. (a) Synchrotron in-plane grazing-incidence (IP-GiD) diffraction indexed to the polar orthorhombic phase ($Pca2_1$) associated with ferroelectric order in HfO_2-based systems. (b) Piezoresponse force microscopy (PFM) phase-contrast images demonstrating stable, bipolar, remanent polarization states that can be overwritten into opposite polarization states.

I_D or gate voltage (Fig. 4.1(b).) Note that, according to classical electrostatics, a typical dielectric at the gate cannot improve the sub-threshold swing with increasing current. This is because increasing current with gate voltage is indication of increasing charge, which in turn indicates an increasing semiconductor capacitance, leading to accompanied loss of gate control. The improvement of SS with increasing current is thus unexpected. What we show in the following is that this seemingly unusual behavior can be quantitatively explained using a ferroelectric gate oxide and related non-linearity stemming from the S-shaped P-E relation.

We start with material characterization. The presence of polar phase in the NCFET gate stack is confirmed by the synchrotron in-plane grazing-incidence diffraction (IP-GID) and piezoresponse force microscopy (PFM) (Fig. 4.2.) The synchrotron IP GID allows us to pick up clean signals of the polar orthorhombic phase from these thin films, while it is hard to pinpoint these phases with standard XRD which is conventionally used for such characterizations. Diffraction spectrum is plotted with respect to the Cu K_α wavelength. The presence of the (120) reflection at $\sim 39°$ is forbidden by the higher-symmetry nonpolar tetragonal polymorph ($P4_2/nmc$.) Additionally, the (111) reflection $\sim 29°$ rather than low-symmetry-split reflections $\sim 27°$ and $31°$ eliminates the presence of the bulk-stable nonpolar monoclinic phase ($P2_1/c$.) The spontaneous polarization can be characterized using the standard box-in-a-box measurement of PFM [29] at large voltages (± 5V) that are sufficient to overcome the depolarization field that locks the NC state in a FE-DE combination [12]. Clear polarization switching in both directions can be seen, providing a further proof of existence of the polar phase.

Anomalous Near-Threshold Behaviours

With the polar phase confirmed, in Fig. 4.3 we look at the I-V for different L_G. While the applied gate bias ranges from -0.4V to 1V in order to avoid stressing or switching the gate stack, it is observed that for all L_G at V_{DS}=0.05V, the NCFET SS is similar to SS of the Control of the same geometry at low drain current (I_D=1pA/μm) but is lowered as the gate bias increases. The I_D at which NCFET sub-threshold swing reaches its minimum is orders of magnitude larger than the OFF current. In addition, the difference in the sub-threshold swing between Control and NCFET devices increases obviously from L_G=100nm to L_G=50nm but decreases from L_G=50nm to L_G=30nm (Fig. 4.5.) Remarkably, these two trends of the NCFETs cannot be explained by assuming a "higher-κ" dielectric gate stack as discussed below.

In the non-degenerate limit, SS can be semi-classically modeled (Fig. 4.1(c)) as

$$SS \approx \left(1 + \frac{C_s + C_d + C_Q}{C_g}\right) \times 60mV/dec \tag{4.1}$$

where the electrostatic coupling between the top of the energy barrier and the source and the drain (respectively C_s and C_d) strongly depends on the device geometry but are relatively independent of the gate stack, the gate capacitance (C_g) is roughly a constant determined by

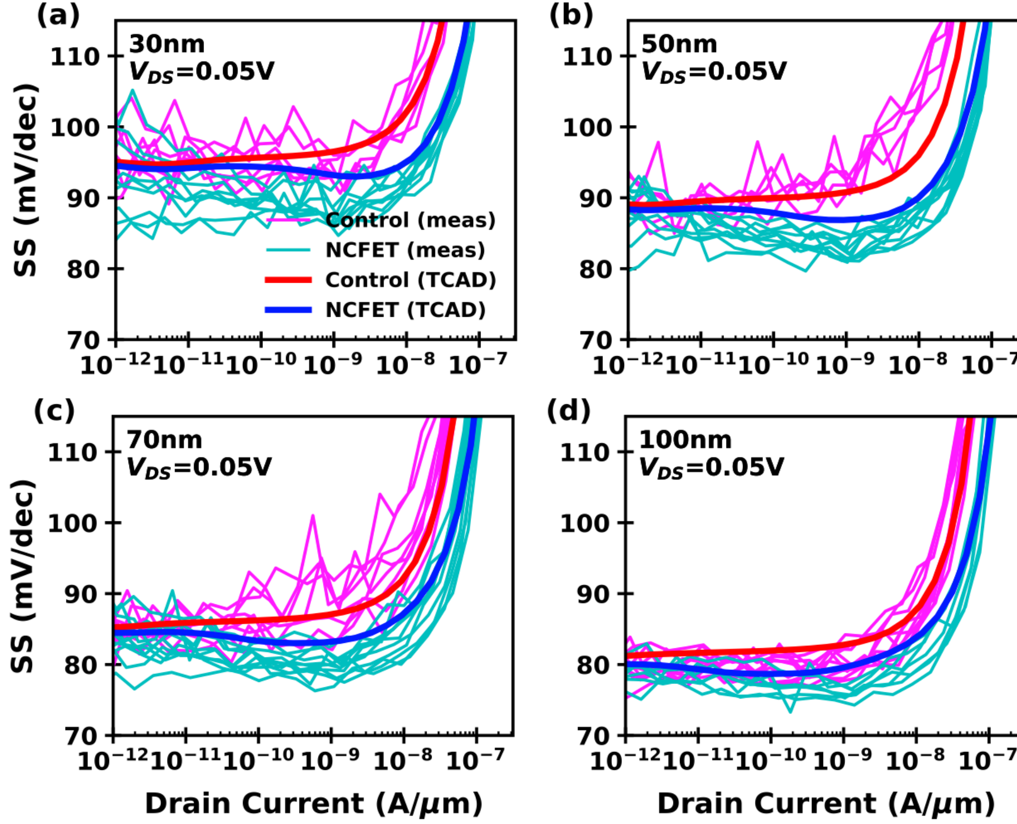


Figure 4.3: Measured and TCAD-simulated SS-I_D relations for Control and NC n-MOSFETs of four gate lengths at V_{DS} =50mV. Applied V_{gs} for measurement is -0.4V ∼ 1V. Multiple devices are measured for each gate length, all of which exhibit negligible hysteresis.

the gate stack equivalent oxide thickness (EOT), and the quantum capacitance (C_Q) depends on the Fermi level in the channel. When V_{gs} is increased, the conduction band energy shift increases both I_D and C_Q, which leads to increasing SS with increasing I_D according to (4.1.)

For two devices of the same geometry at the same drain current, we have

$$SS_1 - SS_2 \approx (C_s + C_d + C_Q) \times (1/C_{g1} - 1/C_{g2}) \times 60mV/dec \tag{4.2}$$

Classically, this difference is dominated by the drastic increase of C_s and C_d as L_g shrinks, resulting in monotonically increasing improvement of SS. The monotonic scaling trends are also seen in simulation works that assume nearly linear ferroelectric P-E relation for NCFET operation voltages [77], which gives rise to NCFET behaviors that can be largely explained by a reduced gate stack EOT. The non-monotonic SS improvement due to NCFET is not explainable with a better interface quality than the Control, either, as it would result in a near-constant C_Q reduction and therefore SS reduction over different L_G. In contrast with normal constant-κ dielectric, these anomalies indicate that NCFET gate capacitance (C_g)

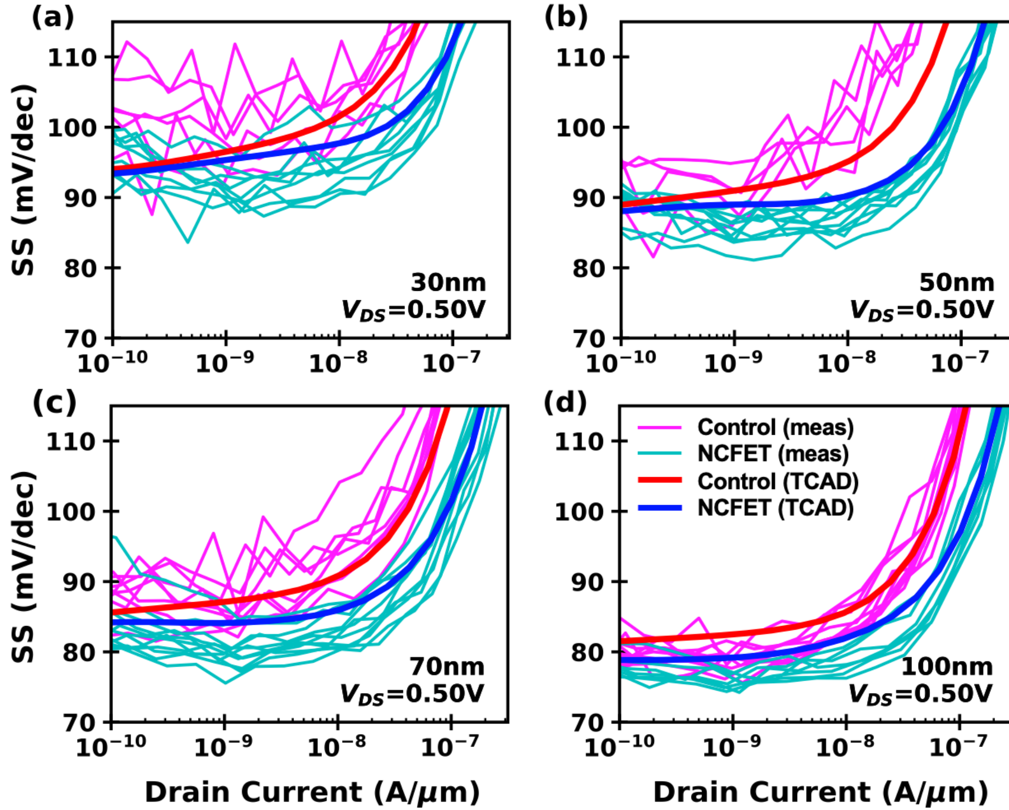


Figure 4.4: Measured and TCAD-simulated SS-I_D relations for Control and NC n-MOSFETs of four gate lengths at V_{DS} =500mV. Applied V_{gs} for measurement is -0.4V ∼ 1V. Multiple devices are measured for each gate length, all of which exhibit negligible hysteresis.

(Fig. 4.1(c)) is affected significantly by both V_{GS} and L_G, of which the explanation lies in the non-linear polarization response to the applied electric field for the NCFET gate stack layer [12, 22], i.e. electric field-induced permittivity enhancement.

4.3 TCAD Model

Sentaurus TCAD [25] simulator parameters are calibrated to C-V (Fig. 4.6) and I-V data through the procedure in Fig. 1.5. The C-V of the gate stacks is first calibrated to an 1-D model to extract the EOT for both the Control and the NC gate stacks. HfO_2 MOSCAP accumulation C-V measurements can be well matched by a 1.1 nm EOT gate stack in the TCAD model. The NC gate stack demonstrates a higher capacitance which could be naively fitted with very high dielectric constant for the HZO stack. However, this hypothetical "super-high-κ" dielectric would not be able to replicate the anomalous sub-threshold behaviors that we observed. Instead, the capacitance boost is captured by the potential amplification effect

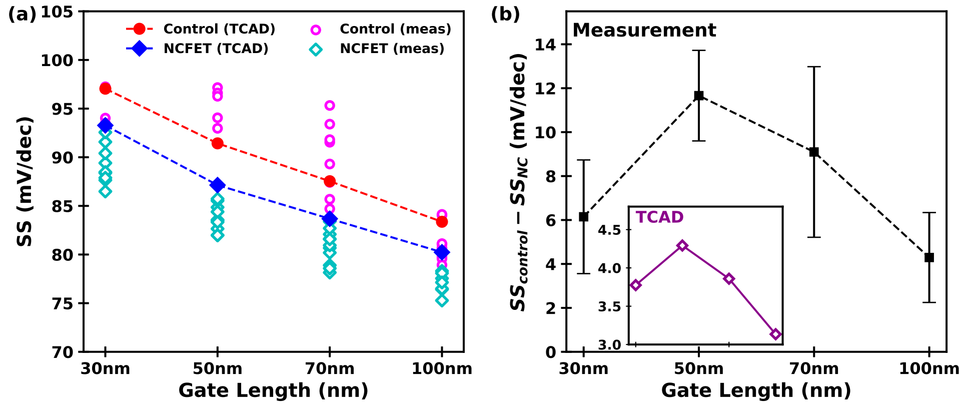


Figure 4.5: (a) SS averaged over I_D=0.1∼1nA/μm at V_{DS}=0.05V for measured and TCAD-simulated Control and NC MOSFET. Each marker presents one device having the corresponding gate length. (b) Experimentally estimated SS improvement for different gate lengths. Each error bar presents the estimated mean and one standard deviation. (Inset) TCAD-modeled Improvements for the four gate lengths.

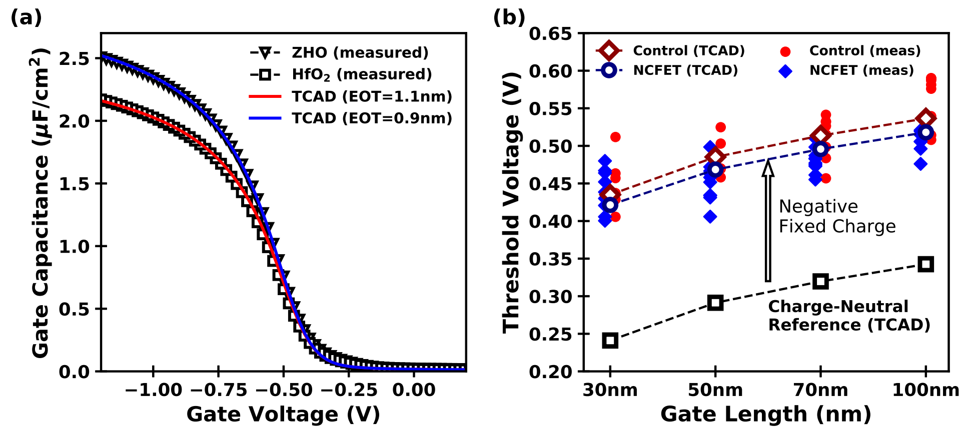


Figure 4.6: (a) Measured and TCAD-simulated C-V for p-type doped MOS capacitors with HfO_2 and ZHO gate stacks. No frequency dispersion is observed up to 300kHz for measurement gate bias range. (b) Constant-current (I_D=10nA/μm) threshold voltage (V_t) at V_{DS}=0.05V for measured and TCAD-simulated Control and NC MOSFET. Each marker presents an extraction for one device. Experimental V_t are 0.18V higher than reference simulations with nearly neutral net oxide and interface trapped charge. The discrepancy is resolved when -0.67μC/cm^2 fixed charge is added to the model.

[9, 12, 22] induced by the polar gate stack in the NCFET.

NC Gate Stack Model

NCFETs are simulated by introducing the Landau-Khalatnikov (LK) model for simulating the gate stack with all non-gate-stack parameters unchanged from the Control devices. The LK parameters are chosen such that the ferroelectric in the positive capacitance (PC) region results in the same gate stack EOT (1.1 nm) as the Control. Therefore, the SS in the low I_D (1∼10pA/μm) region are matched between Control and NCFET devices. As the gate bias is increased, the ferroelectric enters into the NC region, which results in a reduced gate stack EOT of 0.9nm in accordance with the NC stack C-V. This transition from positive to negative capacitance regime explains the SS reduction (Fig. 4.3) with respect to V_{GS} or I_D as well as the non-monotonic SS improvement trend with respect to L_G (Fig. 4.5(b).) The mechanism is discussed in more details in the next section. Note that, although the non-monotonic trend is captured, the SS improvement in (Fig. 4.5(b)) is systematically underestimated by the TCAD model because of unmodelled sources of inter-wafer and inter-device variations. The fitting can be further improved if the constraint of using EOT from C-V measurement is lifted, but we present the constrained fitting which already gives quantitative match to the experimental ranges, demonstrating the consistency between the MOSCAP and MOSFET results.

Threshold Voltage and Fixed Charges

Fig. 4.6 shows measured transistor threshold voltage (V_t) and TCAD calibration results with tungsten work-function set to 4.6eV as extracted from MOSCAP C-V data. Note that Control MOSFET V_t would be underestimated by 0.18V if the effective gate work-function is assumed to be the same as that of the MOSCAPs. We ascribe this discrepancy between the MOSCAP and MOSFET wafers to transistor-specific fabrication processes (such as interlayer dielectric deposition, deposition of via-metals) that include both thermal and plasma processes, which potentially results in a combination of extra fixed charges, or change of the metal work function. A fixed charge density of -0.67μC/cm^2 at the Si/SiO_2 interface can not only account for the V_t shift but also explain the small measured V_t difference between Control and NC devices. The voltage drop on the gate stack induced by the charges is less for the NCFET than the Control because of its larger C_g, which results in a lower V_t that is consistent with the experiments (Fig. 4.6(b).) Note that the introduction of negative fixed charges into the model is consistent with the work as described in Chapter 3 where the transistor fabrication process was the same as used in this study. Notably, the difference in V_t between control and NCFETs could also be explained by change in the interface dipole [78, 79], as the high-κ oxides are slightly different. This would result in a flat V_t shift for each wafer across all L_g. However, as the short channel effects for both wafers are captured by the TCAD model, these constant shifts would not change the main conclusion of this work. The resultant TCAD model captures both the SS-ID and the I_D-V_G characteristics

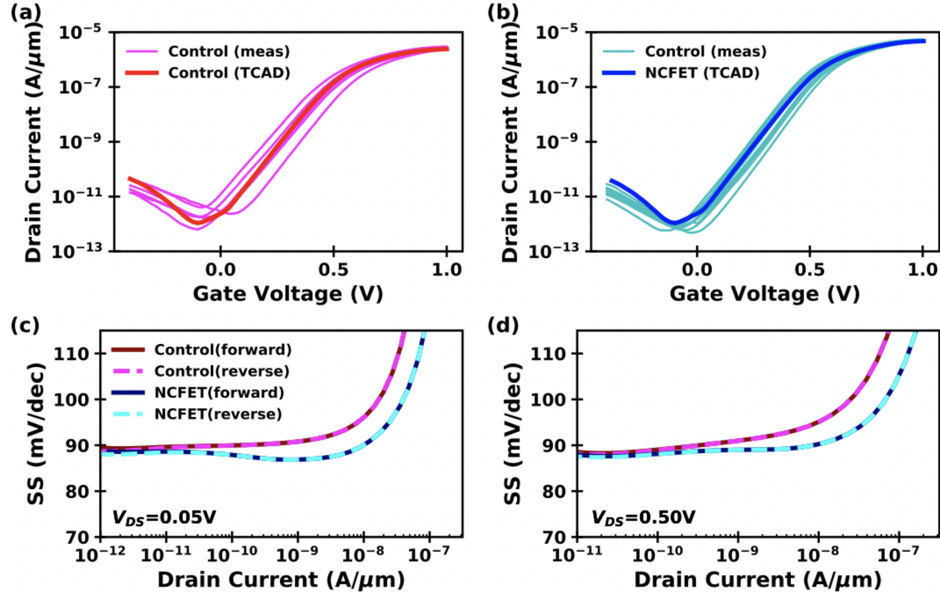


Figure 4.7: (a)-(b) Measured and TCAD-simulated I_D-V_G relations for Control and NC n-MOSFETs with L_G=30nm at V_{DS}=0.5V. Each measurement curve characterizes one MOSFET. (c)-(d) Simulated SS-ID from transient gate-sweep TCAD simulations (L_G=50nm.) The forward and reverse gate voltage sweeps produce overlapping curves, implying zero hysteresis.

(Fig. 4.7(a)(b).) Importantly, the transient TCAD simulations for V_{GS} sweep demonstrates zero hysteresis (Fig. 4.7(c)(d)) in agreement with the experimental results and the concept of NC stabilization.

4.4 Discussion

Nonlinear PC-NC Transitions

Fig. 4.8 and 4.9 show the transitions from positive to negative capacitance region of the NC gate stack. When V_{GS} ramps up, both I_D and the external electric field on the gate stack (E_{ext}=Q_G/ϵ_0) increases. The NC stack polarization starts in the positive capacitance region. As the device enters the NC region, voltage amplification ensues and drives down the sub-threshold swing (Fig. 4.8.) This explains why, for the NCFET, the sub-threshold swing can become steeper as V_{GS} and I_D increases. To understand the observed anomalous L_G scaling trend, the inner fringing field discussed in Chap. 3 again plays a crucial role in the capacitance matching effect. This field leads to the gate stack-to-source and gate stack-to-drain capacitive coupling, which, in turn, reduces Q_G to balance the positively ionized source and drain donor charges for n-MOSFET in the sub-threshold regime (Fig. 3.3.) Therefore,

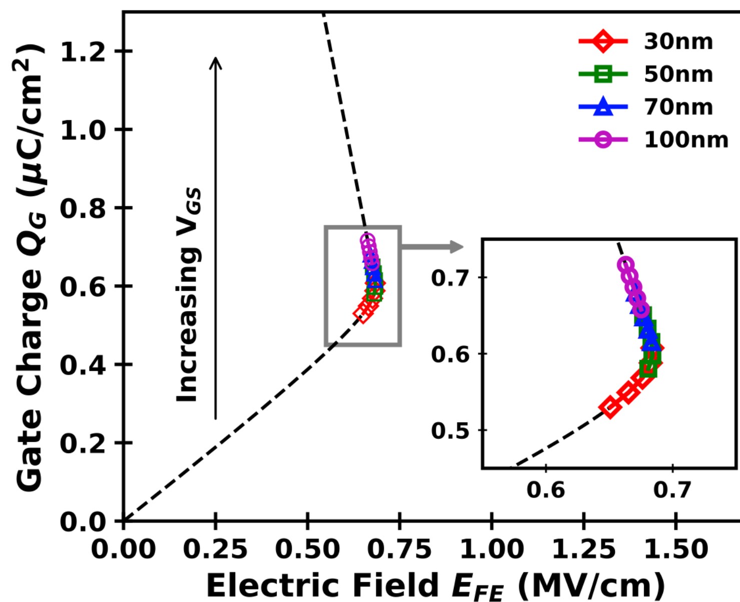


Figure 4.8: Extracted relations between external charge and net electric field for the polar layer characterized at the mid-channel of NCFETs at V_{DS}=50mV. For each gate length, the gate voltage is ramped at which I_D is from 0.1nA/μm to 1nA/μm.

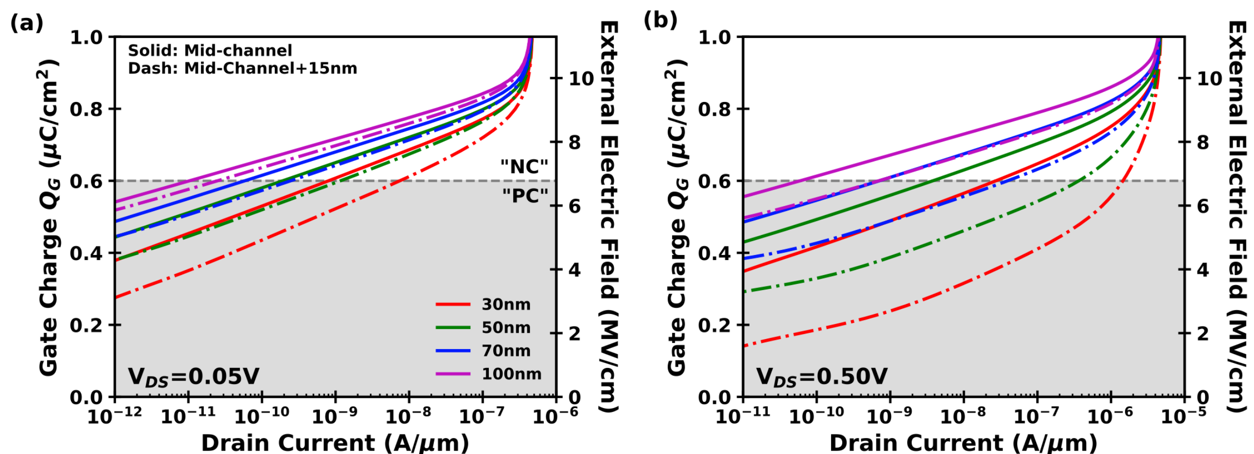


Figure 4.9: Simulated relations between drain current and (left axis) local gate charge density / (right axis) local external electric field at (a) V_{DS}=0.05V and (b) V_{DS}=0.50V for NCFETs. For each gate length, the solid line is extracted at mid-channel gate stack region, and the dash line is extracted at the region 15nm laterally from mid-channel toward the drain. The white-background region corresponds to bias conditions at which the polar layer exhibits negative capacitance, while the dark background stands for positive-capacitance regime.

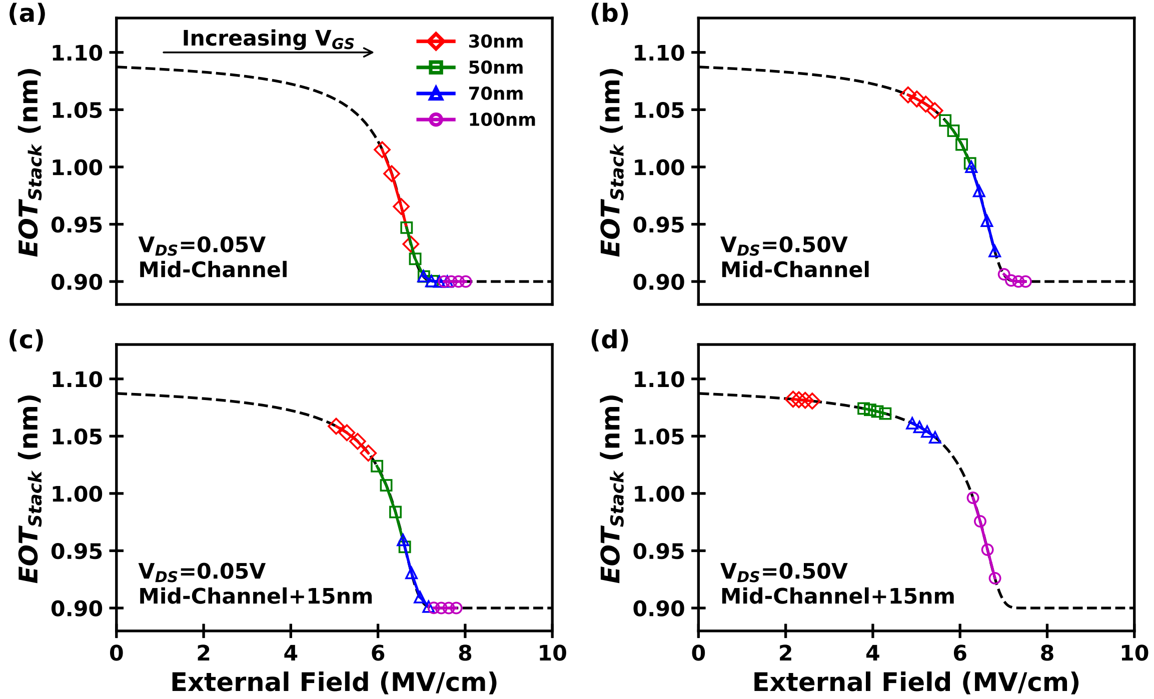


Figure 4.10: Relations between small-signal capacitance-equivalent SiO_2 thickness of the gate stack and applied external electric field extracted at I_D=0.1∼1nA/μm. Extractions are made at different lateral positions under two drain biases.

the external field on the gate stack is a function of lateral position, bias voltages, and L_G (Fig. 4.9.) Each position in each L_G device, at given V_{GS} and V_{DS}, experiences a different external field (Fig. 3.5) and therefore a different electric-field enhanced EOT reduction as shown in Fig. 4.10 and the SS reduction in Fig. 4.3 for small V_{DS}.

The model also explains the absence of near-threshold swing reduction in V_D = 500mV measurements. For V_{DS}=0.50V, the external field becomes a strong function of lateral position, and only a small portion of gate stack passes the nonlinear regime at a given bias (Fig. 4.10(b) and (d).) As a result, a weaker V_{GS} and I_D dependence of SS is observed in Fig. 4.4 than in the V_{DS}=0.05V cases.

Explanations to SS Scaling Trends

The non-monotonic SS improvement in Fig. 4.5(b) can be explained with Fig. 4.8 and 4.10. Although the mid-channel gate stack is in the negative capacitance regime for L_G from 50nm to 100nm at I_D=1nA/μm, it remains in the positive capacitance regime in the 30nm case (Fig. 4.10(a)-(b)) because of the strong inner fringing field. Therefore, the improvement in the sub-threshold swing for NCFETs drops for L_G=30nm as shown in Fig. 4.5(b).

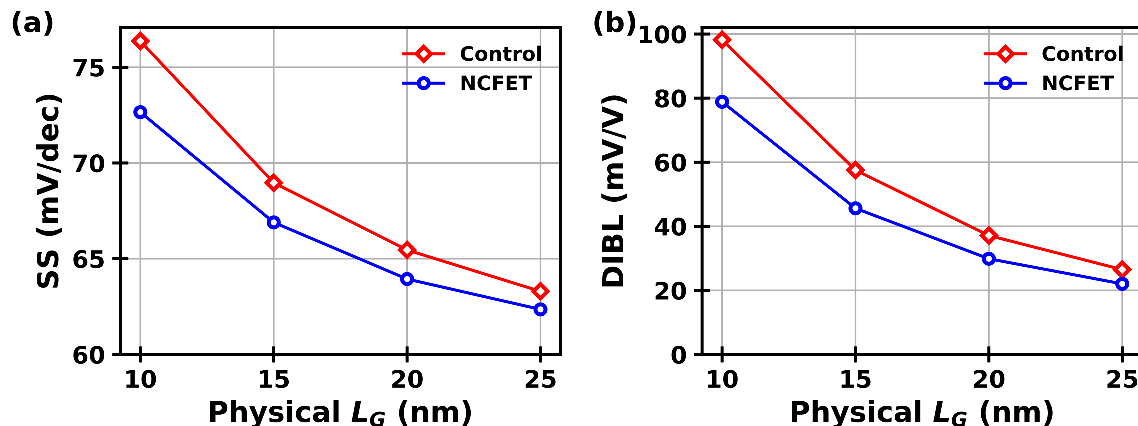


Figure 4.11: Simulated (a) SS and (b) DIBL for various L_G with the Control gate stack and a charge-engineered gate stack that exhibits sustained NC behavior. FinFET structures with 7nm fin thickness are assumed according to the IRDS “5nm” node.

Charge-Shifted S-Curve

Notably, the extracted Q_G at which the NC gate stack starts to exhibit negative capacitance is approximately $0.6\mu C/cm^2$ (Fig. 4.8-4.10.) In other words, the ‘S’ curve is shifted up in the charge axis. The non-zero-centered S curve resembles the P-E relation of antiferroelectric phases that possibly exist in the HZO system. In addition, a large HZO/SiO_2 interface trap density may also lead to such behavior for a ferroelectric system. In this case, due to finite leakage through the IL, the trapped charge at FE/IL interfaces effectively screens the depolarization field that could stabilize the NC state, and FE layer is expected to be outside the NC regime in equilibrium [80]. The mechanism effectively shifts the S-curve when viewed from outside. A similar shift has also been recently reported on the hysteresis curve of Ferroelectric FET memory devices [81]. The difference between the extracted electrical responses of the polar layers in this work and the previous chapter is possibly due to different ALD recipes and cycling as well as the underlying SiO_2 thicknesses which results in different HZO film properties.

Projections With the NC Gate Stack

Finally, we have used our calibrated model to investigate ultra-scaled FinFETs for several L_G options (Fig. 4.11), using the IRDS “5nm” node parameters including 7nm-thick fin [82]. The NCFET with the NC stack presented here and optimized fixed charges to enable sustained NC effect in the operation regime can provide a 2∼3nm (about 15%) reduction of the physical L_G from the Control while achieving the same SS and DIBL. For our calibrated gate stack model, this is achieved by adding approximately $0.65\mu C/cm^2$ or more negative fixed charges to the HZO/SiO_2 interface to push the HZO bias points upward in the gate charge and the

external electric field axes, so that the device is in the NC region across all simulated bias points and positions in the channel. There are multiple ways to control the fixed charge induced shift, especially through gate work function metal engineering exploiting long-range interface dipoles such as those described in [79]. Similar NC benefit to ultra-scaled GAA is expected.

Note that the amount of charge shifts discussed in this section is based on the model which assumes -0.67μC/cm^2 fixed charges to solely explain the MOSFET V_t shift. Possible existence of dipoles may change the number of fixed charges that is extracted from the experiments and therefore the extracted charge shift of the S-curve. Apart from the actual amount of charge quantities, the main conclusions and the drawn physical insights should remain the same.

4.5 Conclusion

To conclude, we developed an FE model that simultaneously explains

1. SS getting steeper with increasing I_D in the sub-threshold regime, and

2. non-monotonic L_G dependence of SS improvement.

These trends cannot be explained by a classical "high-κ" scaling theory. The results show that capacitance matching leads to observation 1, while the inner fringe field induces change in the capacitance matching in short channel MOSFETs and leads to observation 2.

Discussion on NCFET has often focused on observation of $<$ kT/q sub-threshold swing. On the other hand, the S-shaped nonlinearity that leads to negative capacitance could manifest in multiple ways. In fact, reduction in the required supply voltage, without giving up the I_{ON}, depends substantially on a good capacitance matching near the threshold, by which a substantial voltage reduction can be achieved even without going below the kT/q limit in the sub-threshold region. The present work shows such an example of a near-threshold capacitance matching. A quantitative model of the gate oxide opens up new opportunities to exploit the NC effect for reduction in V_{DD} and/or L_G with appropriate device design.

Chapter 5

Performance Modeling of Quasi-Ballistic NCFET

5.1 Introduction

Besides the enhanced gate control at the sub-threshold region which benefits the OFF current, the negative capacitance effect also promises improved gate capacitance in the ON state. [83] has demonstrated 7Å gate stack EOT, lower than the physical thickness of the IL layer, with laminated Hf-Zr-O gate stack deposited through CMOS-compatible atomic-layer deposition (ALD) process. The large capacitance is observed in C-V at 1e13 mobile charge in the channel, sufficient for both logic and RF applications. Besides, no significant mobility degradation as compared with HfO2 control gate stack was observed. Recently, RF characterizations on are reported on 90nm n-type SOI transistors with similar gate stack processing. The intrinsic cutoff frequency (f_{ti}) falls on the conventional scaling trend, indicating that the electron injection velocity agrees with conventional CMOS technologies.

Having a similar injection velocity to conventional devices, NCFET with enhanced intrinsic gate capacitance enables reduced operation voltage. Moreover, the increased gate capacitance and hence the increased transconductance (g_m) amortize the slowdown due to parasitic capacitance that is extrinsic to the channel, benefiting both logic and RF applications. Such benefits can be characterized by increased extrinsic cutoff frequency (f_t) due to the NC gate stack. Based on the aforementioned experimental results, the work in this chapter investigates the opportunity of applying NC gate stack to state-of-the-art transistor processing technology. Careful assessments on the performance of scaled NCFET are carried out through a combination of:

- Practical device structures from process simulations with well-established processing steps and thermal cycle for advanced MOSFET fabrication.
- Self-consistent electrical simulations with gate stack model calibrated to NC stack experiments.

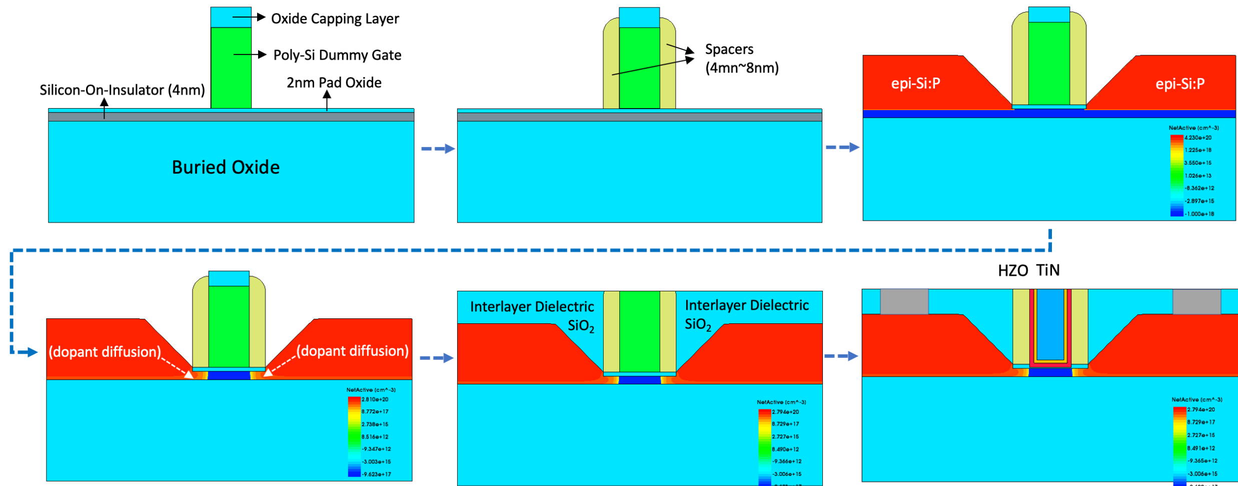


Figure 5.1: Overview of the front-end process flow in this study. The shown example corresponds to the "faceted RSD" flow as described in [85], but is gate-last to accommodate the low thermal budget of the ferroelectric-metal gate stack.

- Rigorous transport models for the quasi-ballistic channel simulation, with carrier scattering rate calibrated to the long-channel experimental results.

The essential purpose of the study is to get a realistic prediction if the gate stack realized in [83, 84] is integrated into available advanced technology. Given the physical parameters, we explore the device design space in order to achieve a high cutoff frequency, which should shed lights on future NCFET designs. The essential process steps taking thermal budgets into consideration is described in the next section, followed by the device modeling methodology and the projected performance.

5.2 Device Structure

In the interest of high-speed operations, potentially enabling next-generation RF applications, self-aligned SOI devices with epitaxially raised source and drain (RSD) are considered. Technologies optimized for RF performance has been productionized, of which one notable example is the 22nm FDSOI technology [86]. The assumed process in this work differs from current SOI technology in that a replacement gate process is selected to accommodate the thermal budget of the HZH gate stack. Fig. 5.1 outlines the simulated front-end process flow. In the temporal order, the process steps and their relevance to device optimization for are described below.

- First, the SOI wafer is prepared with the silicon layer thinned down to the desired body thickness (4nm.) A 2nm padding SiO_2 layer (padox) is uniformly deposited to

protect the silicon from the later polysilicon gate and nitride processing.

- A polysilicon dummy gate stack is then deposited and lithographically defined, where the feature size (the lithographical gate length) is 20nm. Afterwards, Si_3N_4 spacers are formed through isotropic deposition and anisotropic etching, where the deposition thickness is an adjustable parameter for device optimization.

- The padox is removed, followed by surface cleaning and epitaxial Si:P growth. Typically, this can be done at $< 700°C$ [87]. High in-situ doping ($\sim$5e20cm^{-3}) can be achieved to reduce the series resistance and to form good silicide contacts. Typical epitaxial growth would lead to nearly vertical side walls for the RSD. However, the parasitic capacitance between the RSD and the gate can be greatly reduced by introducing "faceted epitaxial" module to achieve a slanted slide wall [85]. Both RSD structures are of interest to this study. For the vertical-wall RSD structure, the RSD height is considered a parameter for device optimization.

- A dopant drive-in anneal is carried out at 800°C. The entire flow is implant-free to prevent silicon amorphization. Therefore, this is the key step to reduce the crowding resistance and to reduce the gate underlap for the spacers. The drive-in time is also a parameter that should be carefully optimized.

- SiO_2 inter-layer dielectric (ILD) are formed through Chemical vapor deposited (CVD), followed by chemical-mechanical polishing (CMP) to achieve a flat surface. The contact hole etching and contact formation are then performed. The contact hole is placed 30nm from the outer edge of the spacer and is 30nm wide. Since low-resistivity contact fabrication usually involves rapid thermal annealing at over 1000°C to enhance the dopant activation and silicidation at over 500°C, the contact module should be placed before the gate stack process which has a thermal budget of 200$\sim$450°C. The detailed silicide contact module is not simulated in this work. However, since the module is not expected to affect the doping profile near the channel, and the subtle change of the device structure near the contact has small effects on the overall device electrical characteristics, the prediction will be reasonable in temrs of performance as long as an appropriate contact resistivity is assumed.

- Finally, the replacement gate stack is formed by etching out the dummy gate and the padox underneath, forming $\sim$8A chemical interfacial SiO_2, and then depositing the NC gate stack through the ALD process described in [83, 84]. The laminated gate metal stack is deposited through ALD and then CVD to achieve appropriate workfunction, followed by CMP to remove excess gate stack materials outside the gate area.

Aside from the slanted side wall RSD structure, the parasitic capacitance can also be reduced by introducing low-permittivity spacers. One notable example for such device optimization is the air-gap spacer process which reduce the spacer permittivity to nearly unity

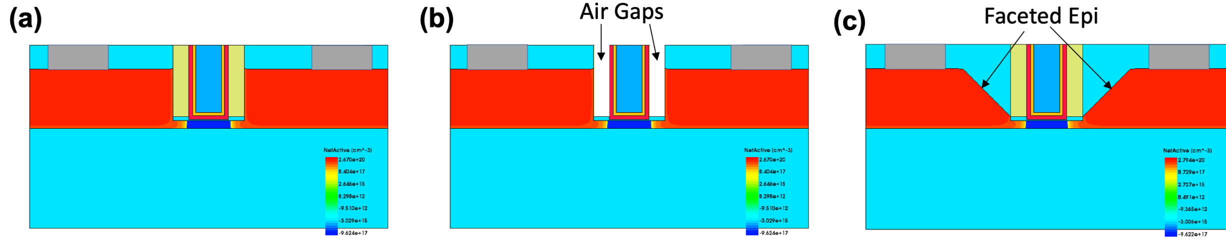


Figure 5.2: Three types of devices studied in this work. (a) The typical RSD structure with nitride spacer and nearly-vertical RSD side walls. (b) The air-gap spacer structure where the nitride spacer is removed from (a). (c) The slanted-side-wall RSD structre formed through specialized epi-Si:P module.

by etching out the nitride spacers after the above process flows [88, 89]. To obtain a comprehensive understanding to the efficacy of different designs which may require additional development efforts, devices with all the three structures in Fig. 5.2 are simulated and analyzed in the following chapters. Note that the padox is not removed for the air-gap spacer structure in this study to ensure that the SOI surface is protected.

5.3 Modeling Methodology

The cutoff frequency (f_t) is the major figure of merit in this study because of its high relevance to both RF and logic applications and its independence to external components such as the gate resistance (r_g) that depends on the layout and the technology. The equivalent small-signal circuit of a MOSFET device (Fig. 5.3(a)) is used to solve for the transistor current gain (H_{21}) through nodal analysis given the values of equivalent circuit elements and an oscillation frequency. Here r_g does not affect the solution of H_{21} and thus can be left unknown. The ft is then obtained by finding the frequency at which $|H_{21}| = 1$.

Definition of Equivalent Circuit Elements

The intrinsic components g_m, r_o, C_{gsi}, and C_{gdi} are defined as the transconductance, output conductance, gate-to-source capacitance, and gate-to-drain capacitance of the channel, respectively. The channel is modeled as the internal region of the device which includes the gate and the spacers and the SOI underneath (the area encompassed by the dashed dark red box in Fig. 5.3(b).) The boundary of the channel (the inner S and inner D) is defined at the junction of the spacer and the RSD. It is defined in such a way instead of at the boundary of the gate and spacer because there can be large lateral and vertical electric field in the area under the spacer which must be solved together with the channel area to obtain the correct electrostatics. Details about the subject is discussed in 5.3. Note that there also exists a

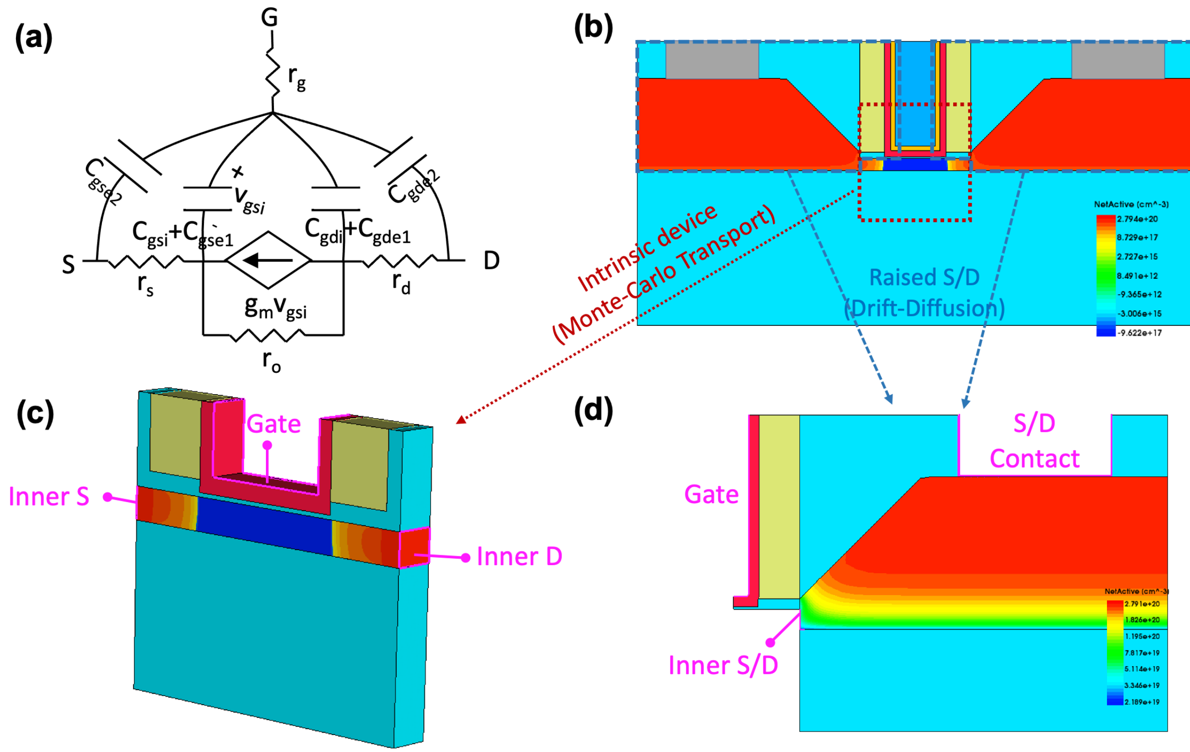


Figure 5.3: (a) The equivalent circuit diagram for modeling small-signal responses of the SOI FET device. (b) Decomposition of the SOI FET into the intrinsic channel and the RSD subsystems. (c) Simulated structure for extracting intrinsic components g_m, r_o, C_{gsi}, and C_{gdi}. (d) Simulated structure for parasitic components associated with the raised source (C_{gse1}, C_{gse2}, and r_s) and the drain (C_{gde1}, C_{gde2}, and r_d.)

finite source-to-drain capacitance for the intrinsic channel, but is negligible when compared to other elements.

Extrinsic to the channel, the RSD can be modeled as a three-port RC network. To model the region accurately without double counting capacitance, the gate, the contact with appropriate interface resistivity, and the RSD must be included in the calculation for related equivalent circuit elements, while the silicon area ascribed to the intrinsic channel must be excluded. An appropriate representation of the RSD subsystem is shown as the area encompassed by the dashed blue polygons in Fig. 5.3(b), where the boundary defining the inner S/D coincides with that for the intrinsic channel. The characteristic frequencies of the RSD subsystems are well above 1THz. Simulations up to 1THz reveal that the gate-to-inner-S/D and the gate-to-external-S/D admittance are purely imaginative, while the inner-to-external-S/D admittance is purely real, all with negligible frequency dispersion. Therefore, the subsystems can be modeled as RC delay lines with lumped elements (C_{gse1}, C_{gse2}, and r_s for the source; C_{gde1}, C_{gde2}, and r_d for the drain) as indicated in Fig. 5.3(a).

Intrinsic Channel Simulation

The L_G of 20nm leads to carrier transport in quasi-ballistic regime. The large electric field in the channel results in rapid changes of spatial and energetic distribution of the carriers. Therefore, it is important to base the prediction of the channel characteristics on fundamental carrier dynamics in conjunction with the silicon band structure and the scattering mechanisms. Valid only when the perturbation from equilibrium statistics is small, mainstream drift-diffusion (DD) solvers in conventional simulators fail to capture the carrier transport in this regime. To resolve the quasi-ballistic transport, the electron transport for the channel is solved with the Synopsis Garand Monte Carlo (MC) Simulator [90]. The tool is capable of simulating transient evolution of semi-classical electron ensembles by resolving the evolution of three-dimensional position and wavevector of each particle. Detailed physical models are described below.

- Contacts are assumed to be connected to equilibrated reservoirs. Particles are injected with Fermi-Dirac distributions according to the Fermi level at the contact.
- Allowed states are described analytically for conduction band electrons. X, L, and G valleys are included.
- The quantum confinement effect is resolved semi-classically through modulation of the band energy according to the solution of the density-gradient equation.
- Self-consistency is achieved by recalculating the electrostatic potential according to the carrier charges every time step (0.05fs), which in turn affects the carrier propagation. As a result, Coulombic interactions between carriers are also resolved.
- Random scattering of the carriers are simulated with scattering rate described by the Fermi's golden rule for various mechanisms such as acoustic phonon, optical phonon, ionized impuriy, and surface roughness scattering. In the event of a single scattering, the momentum and the energy of the scattered particle is modified into the corresponding allowed state. Both intra-valley or inter-valley scattering are possible.

An example of the transfer characteristics simulated with different solvers is provided in Fig. 5.4. The drift-diffusion solver which fails to capture the quasi-ballistic transport would significantly underestimate the drain current. One may also perform a "single-particle" MC simulation which would only consider one carrier propagating in the device for a given simulation instance without self-consistently solving for the electrostatics at each time step. By averaging over many simulation instances, the single-particle MC simulation would produce similar results to the self-consistent ensemble simulation. However, it would slightly overestimate the current because of the lack of long-range Coulombic interaction between carriers which is known to cause momentum transfer between source and channel electrons [91, 92] and reduce the average velocity of channel electrons. To correctly account for the mechanisms discussed above, ensemble MC simulations with self-consistent electrostatics is adopted in this study.

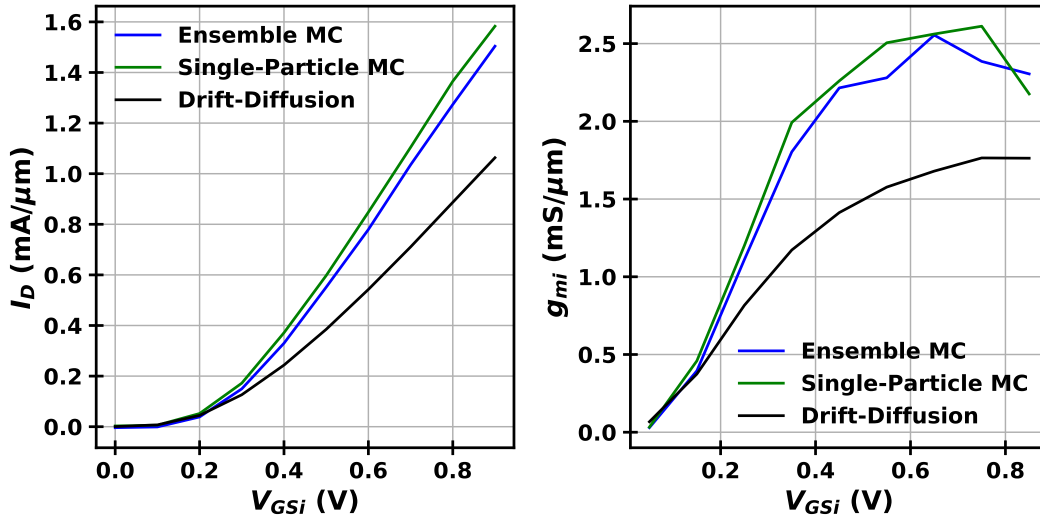


Figure 5.4: Simulated (a) I_D-V_{GS} and (b) g_m-V_{GS} at V_{DS}=0.7V of the same device structure through different transport solvers with similar low-field mobility and optical phonon scattering models. The drift-diffusion simulation produces lower current than the more rigorous solvers. The single-particle MC simulation produces a slightly higher current than the MC simulation with self-consistent electrostatic potential.

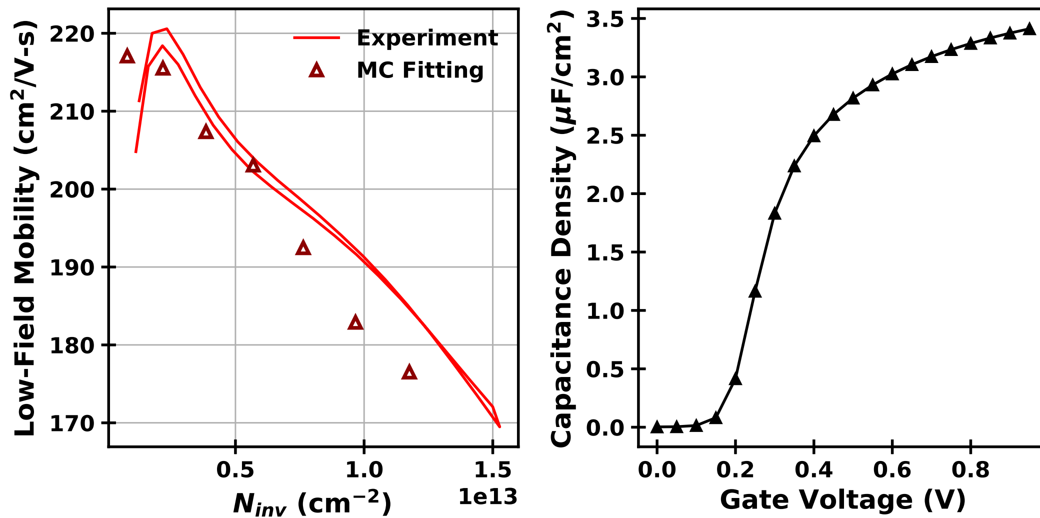


Figure 5.5: (a) The measured and MC-simulated low-field electron mobility for long channel NCFET devices. (b) The simulated long-channel capacitance density.

Simulation setup

The intrinsic channel transport are simulated with the structure in Fig. 5.2(c) through the MC simulator. The overall device structure is cropped for the intrinsic channel area and extruded into the width dimension. Besides, 2nm of highly-doped (2e20cm^{-3}) extension regions are added between the intrinsic channel region and the inner S/D contacts to stabilize the boundary condition for particle injections. Note that the Fermi level at the contact is dynamically determined according to the ensemble distribution, so the high doping of the extension region does not enforce the distribution of carriers injected or significantly affect the carrier profiles in the channel. Without the extra extension region, the solution for contact carrier distribution can be unstable and give rise to unphysical results. Surface orientation of (001) and channel direction in [110] are simulated. The scattering rate is calibrated to the low-field mobility from NCFET experiments [84] in long-channel simulations (Fig. 5.5(a).) The gate workfunction is assumed to be 4.35eV, and the EOT of the gate stack is set to 7Å according to the measurement results in [83], which in conjunction with the quantum confinement effects gives rise to long-channel capacitance density shown in Fig. 5.5(b). For the 4nm-thick silicon body, the geometrical quantum confinement effect would lead to approximately 100meV subband energy separation between the Δ_4 and the Δ_2 valleys. The energy offset is manually added into the analytical band profile and lifts the 6-fold X valley degeneracy.

Estimation of g_{mi} and r_o

For each simulated bias point, an initial solution is obtained from a drift-diffusion solver and used to initialize the MC transient simulation. The MC simulation is carried out for at least 40000 time steps (2ps in the simulation domain) where the initial period of 30000 steps are reserved for carrier redistribution to reach a steady state and thus not included in the electric current estimation. After the initial period, the net flux of electrons for all subsequent time steps are averaged to estimate the electric current. The simulation is considered converged only if the standard deviation of the current estimate is within 1% of the average. The g_{mi} and r_o can hence be calculated through differentiating the steady state results across different bias points.

Estimation of C_{gsi} and C_{gdi}

The MC solver used in this study does not produce capacitance estimates, so the extractions of C_{gsi} and C_{gdi} are based on the drift-diffusion solution. Although failing to capture the transport, the drift-diffusion solver gives plausible description to the electrostatics and therefore should predict the capacitance reasonably. This is confirmed through the similarity between the electron density profiles for the steady-state MC solution and the drift-diffusion solution. The comparison also reveals that the sheet channel carrier density for the drift-diffusion solutions is slightly higher than the MC solutions, which implies that the drift-diffusion solver may slightly overestimate the C_{gsi}. Detailed comparisons are provided

in 5.4. In terms of the performance study, this results in a mild pessimism, but we stress that this is favored over optimism in the interest of reliable benchmarks.

Source and Drain Parasitics Simulations

The RSD-related parasitic elements are simulated with the structure in Fig. 5.3(c) through drift-diffusion solvers. Generally, small electric field in the silicon suffices to drive a large current for the RSD subsystem because of the high doping throughout. Therefore, drift-diffusion equation describes the transport in RSD well for the operation biases of interest. The mobility models accounting for phonon scatterings, Coulombic scattering, surface roughness scattering, and degradation due to Si thickness fluctuations are simulated for electrons. To capture the effect of the ohmic contact resistance, an interface resistance at the external S/D contact is simulated in a spatially distributive manner. The contact resistivity is assumed to be 1.5nΩcm^2 which can be achieved through epitaxial Si:P, laser annealing, and novel silicidation process [93].

For a given bias condition for the intrinsic channel, the gate voltage and the inner S/D voltage is set accordingly, and the external S/D contact voltage is varied to search for the RSD bias that is consistent with the drain current from the channel MC simulation. After the DC bias point is determined, a linearized system analysis is performed to extract the small-signal admittance matrix for the contacts, which can be translated into C_{gse1}, C_{gde2}, and r_s for the raised source and C_{gde1}, C_{gde2}, and r_d for the raised drain according to the discussion in the beginning of this section. Having calculated all the components in Fig. 5.3(a), the ft can then be estimated accordingly.

5.4 Results

For the three types of structures in Fig. 5.2, the dopant diffusion time is varied between 2min and 15min, the spacer thickness is varied between 4nm and 8nm, and the RSD height is varied between 20nm and 30nm. The three structures are optimized individually and then compared. The RSD height has negligible effects on the doping profile of the intrinsic channel including the spacer region, while all the three parameters have considerable effects on the RSD characteristics.

Intrinsic Device

Carrier Dynamics

The MC steady-state averaged carrier profiles of a simulated transistor are shown in Fig. 5.6 and 5.7. Due to the effective potential accounting for the quantum confinement effect, the electron density is lowered near the interfaces (Fig. 5.6(a).) The profiles for carriers traveling in different directions can be summarized separately. The mobile electrons having a positive velocity in the source-to-drain direction and thus contributing positively to I_D are classified

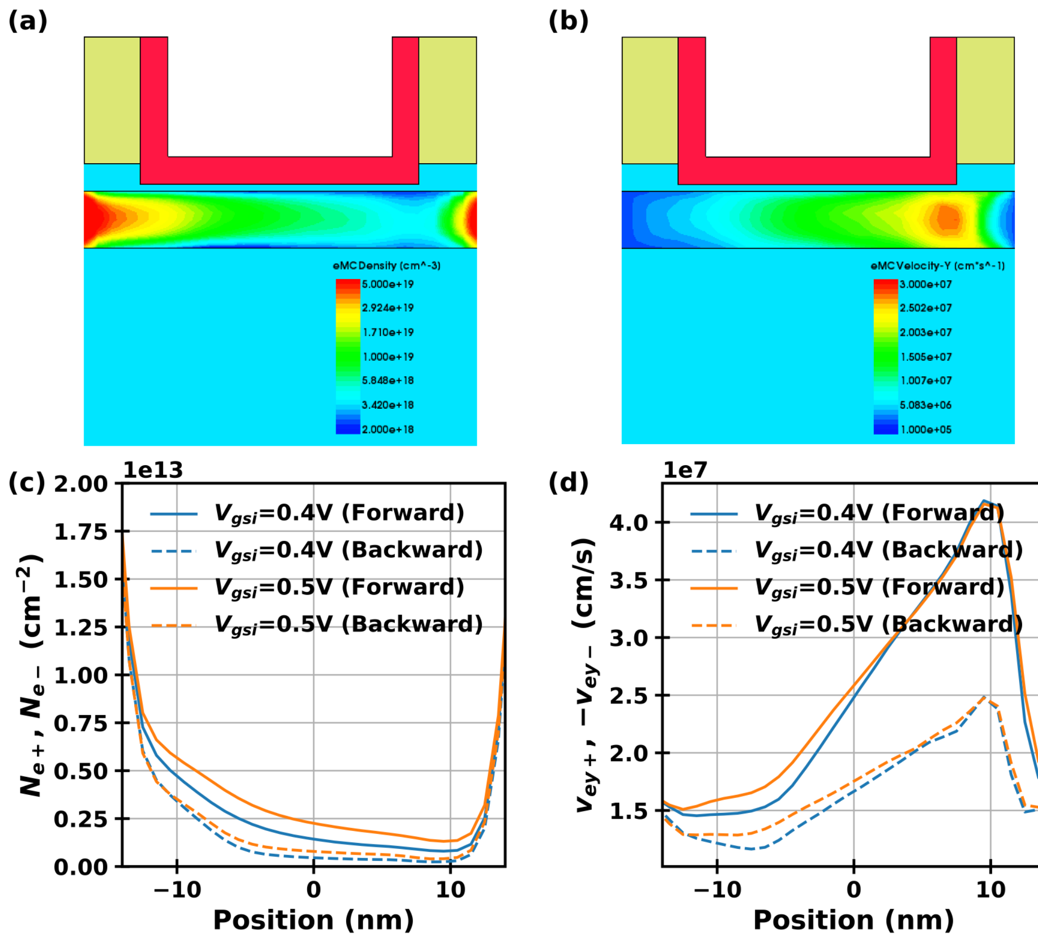


Figure 5.6: MC steady-state profile of (a) conduction electron density and (b) electron velocity in the transport direction for a 4nm-spacer NCFET at $V_{GSi} = 0.45V$ and $V_{DSi} = 0.7V$. (c) The sheet density of the right-going electrons (N_{e+}) and the left-going electrons (N_{e-}.) (d) The averaged velocity in the transport direction for the right-going electrons (v_{ey+}) and the left-going electrons (v_{ey-}.) The x-axes for (c) and (d) represents the lateral position from the inner source (-14nm) to the inner drain (14nm).

as the "right-going" electrons, with local sheet density N_{e+} and averaged velocity in the transport direction v_{ey+}. The left-going electrons are defined vice versa, with sheet density N_{e-} and averaged velocity v_{e-}. Near the source and the drain where the electric field is low, the population for N_{e+} and N_{e-} are similar (Fig. 5.6(c),) close to the classical expectation that is implicit in a drift-diffusion solver. Because of the low electric field, v_{ey+} near the source and v_{ey-} near the drain are expected to be close to the thermal injection velocity, which appears to be around $1.5 \times 10^7 cm/s$ that is consistent with the calculation in [94]. This is larger than the injection velocity in the degenerate limit (about $1.2 \times 10^7 cm/s$) because of the degenerate carrier statistics due to high carrier density and the Δ_4 energy offset stemming from the quantum confinement effect.

The profiles in Fig. 5.6(c) and (d) are signature of quasi-ballistic transport phenomena. As the carriers are injected into the channel, the electric field of the potential barriers gives rise to a force that is opposite to the injection direction. Almost all mobile electrons injected from the drain are reflected because of the large potential barrier formed by the gate-to-drain voltage, leading to very low N_{e-} near the drain edge. In contrast, the gate-to-source voltage forms a smaller potential barrier which partially reflects carriers injected from the source, so N_{e+} does not fully decay. The right-going electrons that are able to pass the top of the potential barrier gets accelerated by the lateral electric field, which leads to increased v_{ey+} as the position increases. Meanwhile, scattering events change some of the right-going electrons into the left-going states, leading to non-zero N_{e-} in the channel. v_{ey-} in the channel is highly correlated to v_{ey+} at the same position because the scattering event does not fully relax the energy of the scattered carriers. When the right-going electrons enter the drain, the acceleration due to the electric field has stopped, while significant optical phonon scattering leads to quick relaxation of the high carrier energy and thus reduced v_{ey+}.

Comparisons with Drift-Diffusion

The quasi-ballistic MC simulations produce charge profiles (N_e) that are close to the DD results (Fig. 5.7(a),) except that the N_e from MC simulations are smaller than the DD results in the regions with high electric field (5nm to 10nm) in the figure. Correspondingly, the gate capacitance from MC is lower than DD in the region (Fig. 5.7(b).) The MC simulation has an increased capacitance density at the drain edge where significant energy and momentum relaxations of the source-injected carriers take place. After integration, the overall C_{gsi} from MC is slightly lower than that from DD, justifying the use of DD simulation for capacitance estimations. The difference in the N_e for the two simulators can be understood through the v_{ey} difference (Fig. 5.7(c).) For DD simulation, the velocity saturation model enforces v_{ey} to be lower than $1 \times 10^7 cm/s$, while the acceleration due to high field can result in non-equilibrium statistics with v_{ey} that is much higher than the thermal injection velocity. As a result, the DD model predicts prematurely the location at which v_{ey} saturates, preventing N_e from decreasing after the velocity saturation point, while in MC simulations N_e keeps decreasing because of the velocity increase.

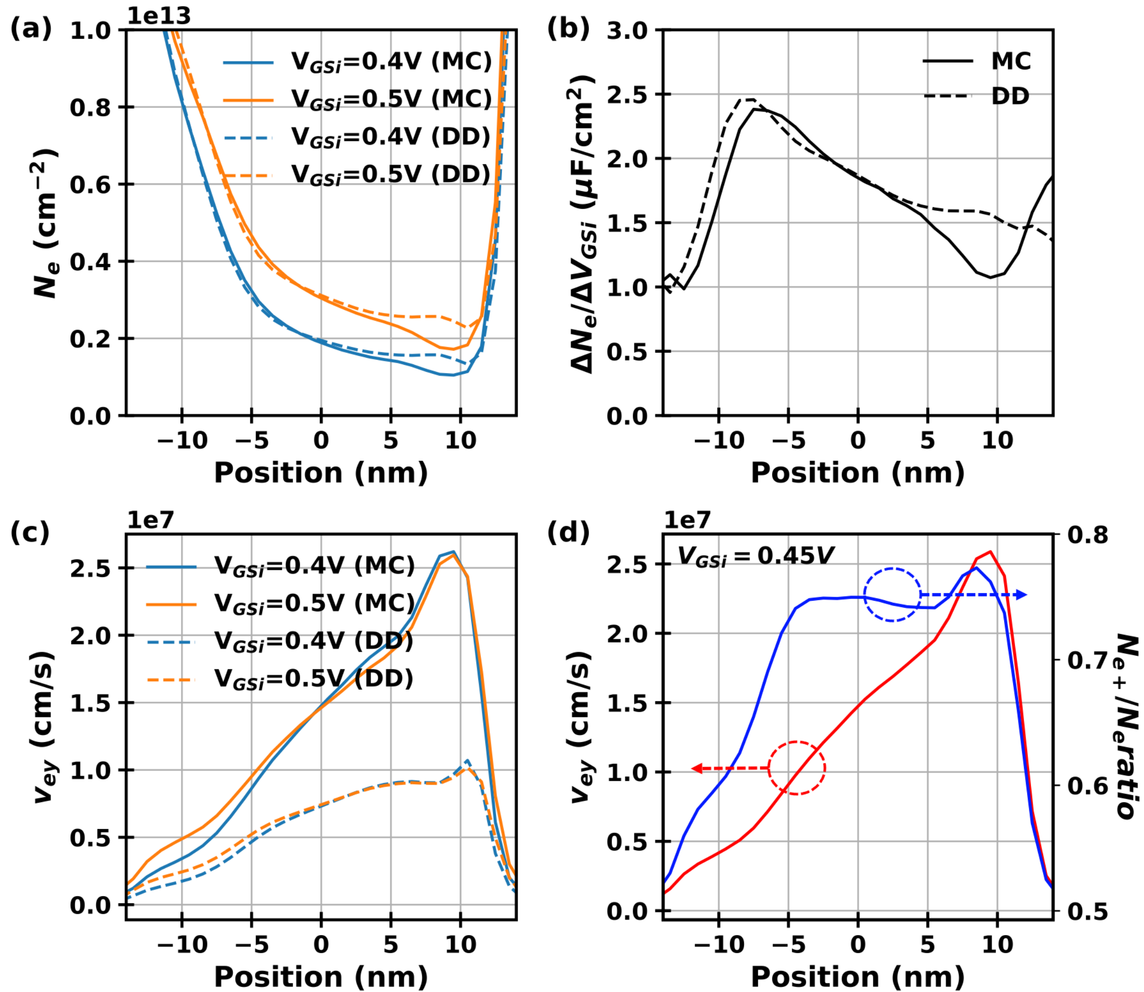


Figure 5.7: (a) The sheet conduction electron density (N_e,) (b) the local capacitance density characterizing the response of the N_e to the change of V_{GSi}, and (c) the averaged electron velocity in the transport direction (v_{ey}) at $V_{GSi} = 0.4V$ and $0.5V$. Results from different transport simulators (DD and MC) are shown. (d) The v_{ey} and the ratio of the N_{e+}/N_e at $V_{GSi} = 0.45V$. The device is the same as that shown in Fig. 5.6, and the position follows the same definition.

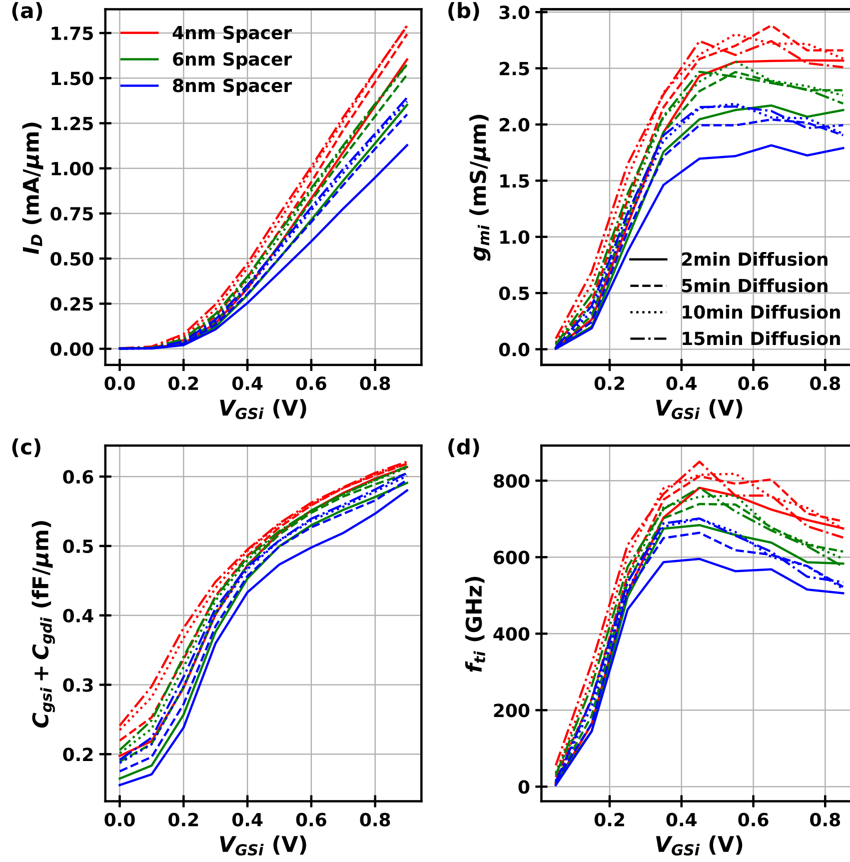


Figure 5.8: Simulated (a) I_D-V_{GSi}, (b) g_{mi}-V_{GSi}, (c) $(C_{gsi} + C_{gdi})$-V_{GSi}, and (d) f_{ti}-V_{GSi} at V_{DSi}=0.7V for faceted RSD devices with various spacer thicknesses and dopant diffusion times. All subplots share the same legends. Small fluctuations of g_{mi} and f_{ti} are due to the random nature of MC simulations.

Operation Bias Voltages

The simulated transfer characteristics (Fig. 5.8) is nearly fully saturated at when V_{DSi}, the voltage between the inner D and S, is 0.7V. For all the simulated structures, the g_{mi} starts to saturate as V_{GSi}, the voltage between the gate and inner S, reaches 0.4V. On the other hand, the C_{gsi} keeps increasing because of redistribution of inversion charge as the gate voltage increases. Classically, this can be understood as the reduction of pinch-off region that is caused by the saturation drain bias. The cutoff frequency of the intrinsic device (f_{ti}) can be estimated as

$$f_{ti} \approx \frac{g_{mi}}{2\pi(C_{gsi} + C_{gdi})} \tag{5.1}$$

which gives maximal values at $V_{GSi} = 0.45V$ for all the simulated cases. Therefore, the operation bias for f_t optimization is determined at $V_{GSi} = 0.45V$ and $V_{DSi} = 0.7V$, which

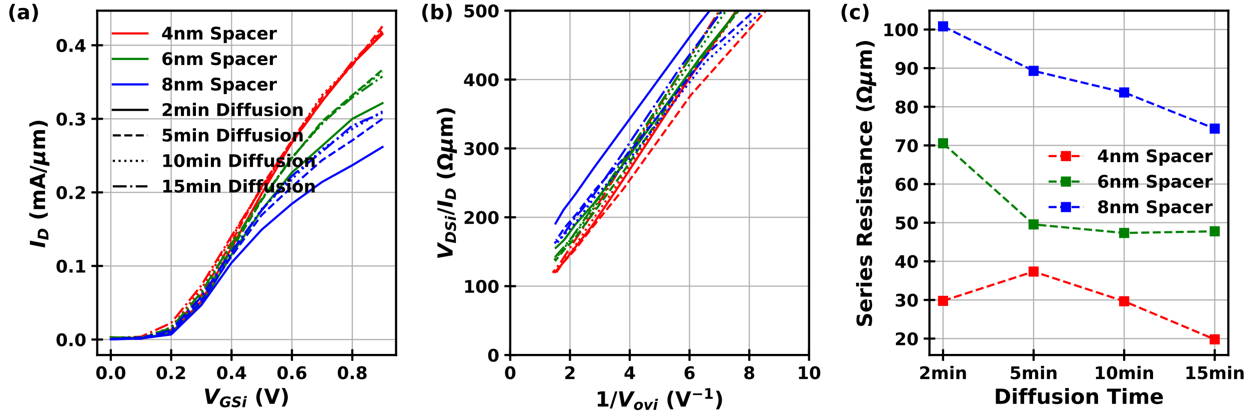


Figure 5.9: Simulated (a) I_D-V_{GSi}, (b) V_{DSi}/I_D-$1/V_{ovi}$, and (c) the extracted series resistance for the intrinsic device at V_{DSi}=0.05V for faceted RSD devices with various spacer thicknesses and dopant diffusion times. The legend for (a) also applies to (b).

satisfies the requirement for low-voltage operations.

Note that the g_{mi} slightly over 2.5mS/μm for 4nm-spacer devices can be explained with the profiles in Fig. 5.7. According to Fig. 5.7(b), the inversion capacitance density (C_{inv}) at the top of the barrier is approximately $2.5\mu F/cm^2$ at $V_{GSi} = 0.45V$. The virtual source velocity can be approximated with Fig. 5.7(d), where the location of the virtual source is defined as the location where the N_{e+}/N_e ratio reaches the first peak (at about -4nm.) The v_{ey} at the virtual source is slightly over $1 \times 10^7 cm/s$. The g_{mi} is expected to be C_{inv} multiplied by the virtual source velocity, which is consistent with the observations in Fig. 5.8(d).

Effect of Extension Regions Under the Spacers

Note that while the symbol assignment in this study suggests "intrinsic", the extractions here include not only the gate-defined channel but also the effects of the extension regions under the spacers, because it is an integral part of the MC simulations. Fig. 5.8(a) and (b) demonstrate the importance of dopant engineering under the spacer region. When the spacer is as wide as 8nm, a significant degradation of the drain current and transconductance from the other cases is observed. When the spacer thickness is small enough (e.g. 4nm,) a short diffusion time suffices to enable good electrostatics under the spacer. The effect is studied in further details through the extraction of series resistance for the spacer regions. MC simulations at a low V_{DSi} (50mV) are performed (Fig. 5.9(a)), and the results are characterized by extrapolating the relationship between the total large-signal resistance (V_{DSi}/I_D) and the inverse of the intrinsic overdrive voltage $1/V_{ovi} = 1/(V_{GSi} - V_{ti})$ where V_{ti} is the threshold voltage extracted through the method in [95] (Fig. 5.9(b).) The series resistance for 8nm-spacer devices is indeed larger than the counterparts, but the $\sim 50\Omega\mu m$ series resistance on each side does not fully explain the g_m degradation that is shown to be

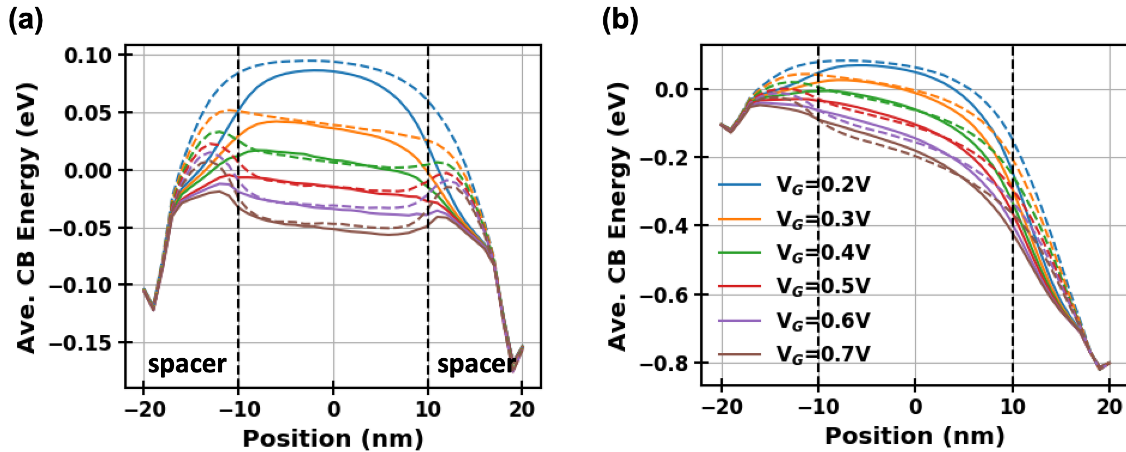


Figure 5.10: The vertically-averaged conduction band energy for the SOI versus the lateral position from inner S to inner D at various V_{GSi} for (a) V_{DSi}=50mV and (b) V_{DSi}=0.7V. Vertical dashed lines indicate the boundary for the gate stack and the spacers. The simulated devices have 8nm spacers undergoing 2min (dashed curves) or 10min (solid curves) dopant diffusion at 800°C. The 2min-diffusion cases have the top of the potential barrier in the spacer extension region at the source side and hence much weaker gate control of the barrier height as compared to the 10min-diffusion cases.

more than 30% for the 8nm spacers.

Instead, the degradation is mainly ascribed to the worsened gate control that can be induced by poorly engineered extension regions for short-channel devices. When the gate underlap is large, the top of the channel potential barrier, which directly affects the current flow, is dragged towards the source (Fig. 5.10.) When the top of the barrier is located under the spacer, it is no longer controlled directly through the gate stack, but an additional series capacitance through the spacer and padox reduces the gate capacitance seen by the top of the potential barrier. Such effect is seen for both low and high V_{DSi} in the 8nm-spacer device simulations without sufficient dopant diffusion (dashed curves in Fig. 5.10(a) and (b).) This is an intrinsic degradation to the channel gate control, and therefore would not be characterized as the series resistance. However, it has strong implications for the short-channel device optimizations to complement the improvements introduced by EOT scaling.

Extractions Summary

At the chosen bias point, the drain current and small signal responses are summarized in Fig. 5.11 and Fig. 5.12. The sub-threshold swing (SS) is also reported according to the drift-diffusion simulations. Note that the drift-diffusion simulations captures the electrostatics correctly when the electrostatic screening effect due to mobile carriers is weak. In general, the

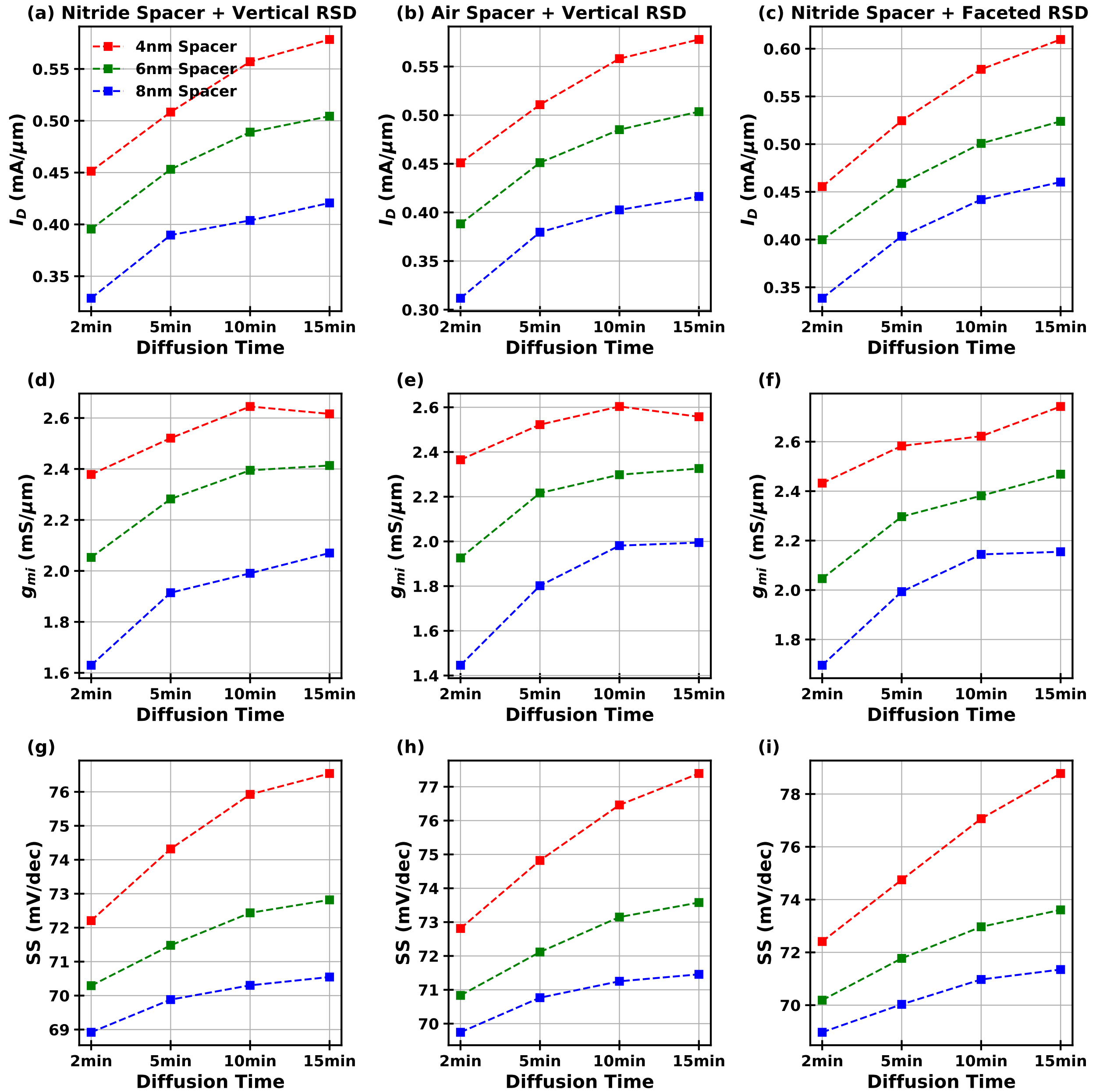


Figure 5.11: Extract (a)-(c) I_D at V_{GSi}=0.45V and V_{DSi}=0.7V, (d)-(f) g_{mi} at at V_{GSi}=0.45V and V_{DSi}=0.7V, and (g)-(i) SS at V_{GSi}=0V and V_{DSi}=0.7V. The same column of subplots represents the same type of device structure as indicated in (a)-(c).

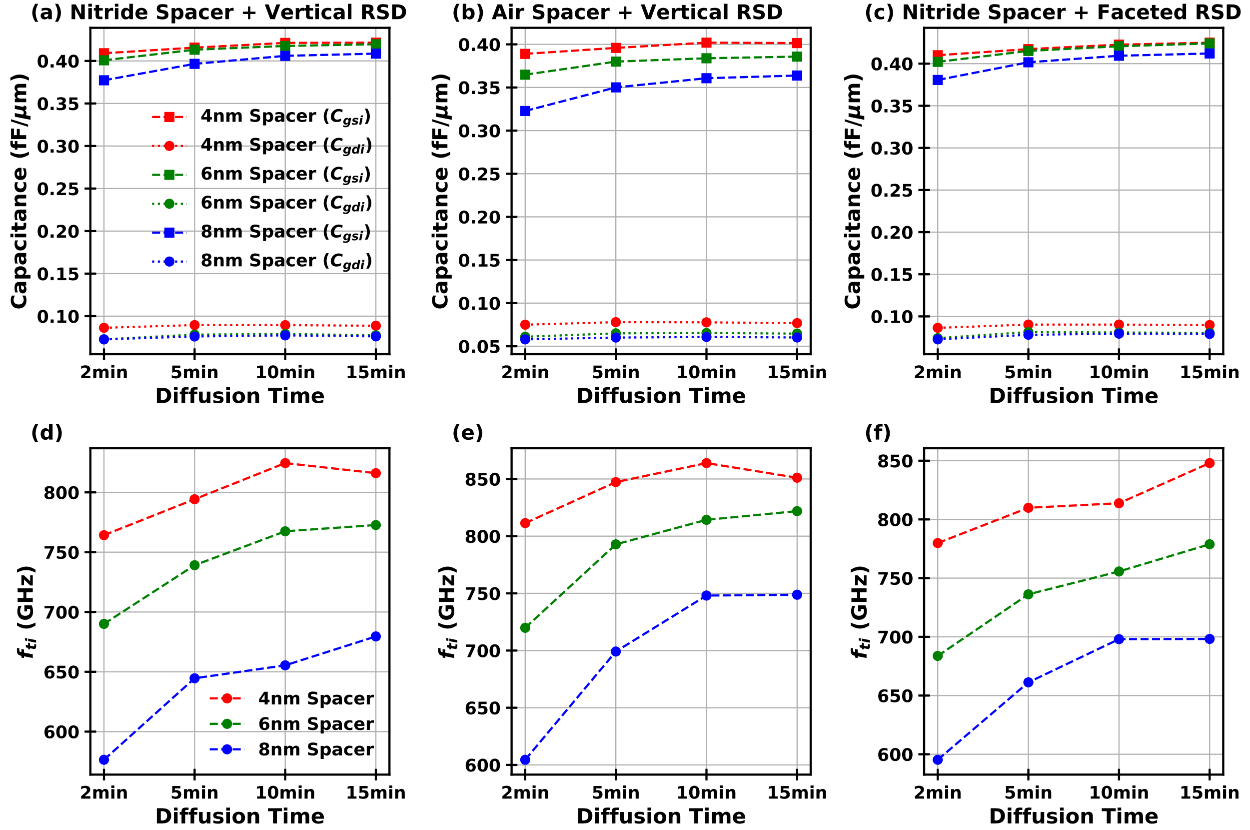


Figure 5.12: Extract (a)-(c) C_{gsi} and C_{gdi} at V_{GSi}=0.45V and V_{DSi}=0.7V, (d)-(f) f_{ti} at at V_{GSi}=0.45V and V_{DSi}=0.7V. The same column of subplots represents the same type of device structure as indicated in (a)-(c).

transfer characteristics follow expected trends with respect to the spacer profiles. Narrowed spacers or extended dopant diffusion times lead to increased I_D, g_{mi}, and SS. Given the same spacer width and diffusion time, the dopant diffusion is slightly deeper for the faceted RSD structure than the other two, resulting in the slightly increased I_D and g_{mi} and degraded SS. As compared to the nitride spacers, the air-gap spacer devices have slightly degraded gate control and hence the I_D, g_{mi} and SS. However, when the spacer extension is doped sufficiently, such degradation is minimal. The r_o (not shown) is approximately 5$\sim$10k$\Omega\mu m$ for the devices, which has small effects on the cutoff frequencies.

The relative changes of capacitance for different conditions are dominated by the outer outer-fringing-field coupling between the gate and the S/D spacer extension regions. This coupling is the strongest near the edge of the gate. Therefore, devices with increased extension doping near the gate edge have increased C_{gsi} and C_{gdi}. Notably, the air-gap spacers effectively reduce the outer fringing field. Although the air-gap spacers have a much larger impact to the RSD parasitics as compared to the intrinsic channel, its benefit can already

be observed in Fig. Fig. 5.12(b) and (e) especially through the increase of f_{ti}.

Raised Source and Drain

Source and Drain Asymmetry

The RSD at the source and the drain sides are simulated separately (Fig. 5.13 and 5.14). The same gate bias leads to slightly different carrier distributions between the source and the drain sides. Because the silicon is biased at around 0V for the source and around 0.7V for the drain, the 0.45V gate bias would induce surface accumulation of electrons for the source-side RSD, while the drain side undergoes a slight surface depletion because $V_{GDi} = -0.25V$. Therefore, for the same device R_s is smaller than R_d (Fig. 5.13 and 5.14, (g)-(i).) Nevertheless, the asymmetry is weak because the outer-fringing field induced by the gate is small, and the abundant RSD carriers screen the field effect effectively.

Extractions Summary

Both the air-gap spacer and the faceted RSD structures significantly reduce the parasitic capacitance (Fig. 5.13 and 5.14, (a)-(f).) For our setting, the air-gap spacer appears to be more effective because it simultaneously enables lower series resistance (Fig. 5.13 and 5.14, (g)-(i)) and lower parasitic capacitance than the faceted RSD structure. Therefore, it is expected to achieve the highest performance among he three structures given that it also provides the highest f_{ti}. Detailed effect of the process parameters will be discussed in the next subsection.

Thanks to the high in-situ doping and the low contact resistivity, the series resistance for the RSD is lower than $50\Omega\mu m$ for all the simulated cases. This corresponds to less than 12% degradation for the DC transconductance ($g_{mi}/(1 + g_{mi}R_s)$) from the intrinsic device. Similarly, the RSD voltage drop to support the channel current in Fig. 5.11(a)-(c) is around 10∼25mV for each side (Fig. 5.13 and 5.14, (j)-(l).) Therefore, the external drain-to-source voltage suggested by the study is generally less than $V_{DSi} + 25mV \times 2 = 0.75V$. Because the quasi-ballistic device is well saturated at $V_{DSi} = 0.7V$, it may be further reduced without significantly affecting the performance.

Summary and Design Perspectives

It is important to balance different equivalent circuit elements to maximize the performance through fine-tuning the device design. Increasing the RSD thickness is beneficial to the series resistance, but comes at a cost of increasing the parasitic capacitance. While reduced spacer thickness is effective for increasing g_{mi} (Fig. 5.11(d)-(f),) it also increases the RSD capacitance (Fig. 5.13 and 5.14, (a)-(f).) To achieve a high g_{mi}, another effective strategy is to use a moderate spacer thickness and to diffuse the dopants for a sufficient time to eliminate source and drain doping underlap. Meanwhile, increased diffusion time reduces the doping concentration in the RSD, which increases the R_s and R_d (Fig. 5.13 and 5.14,

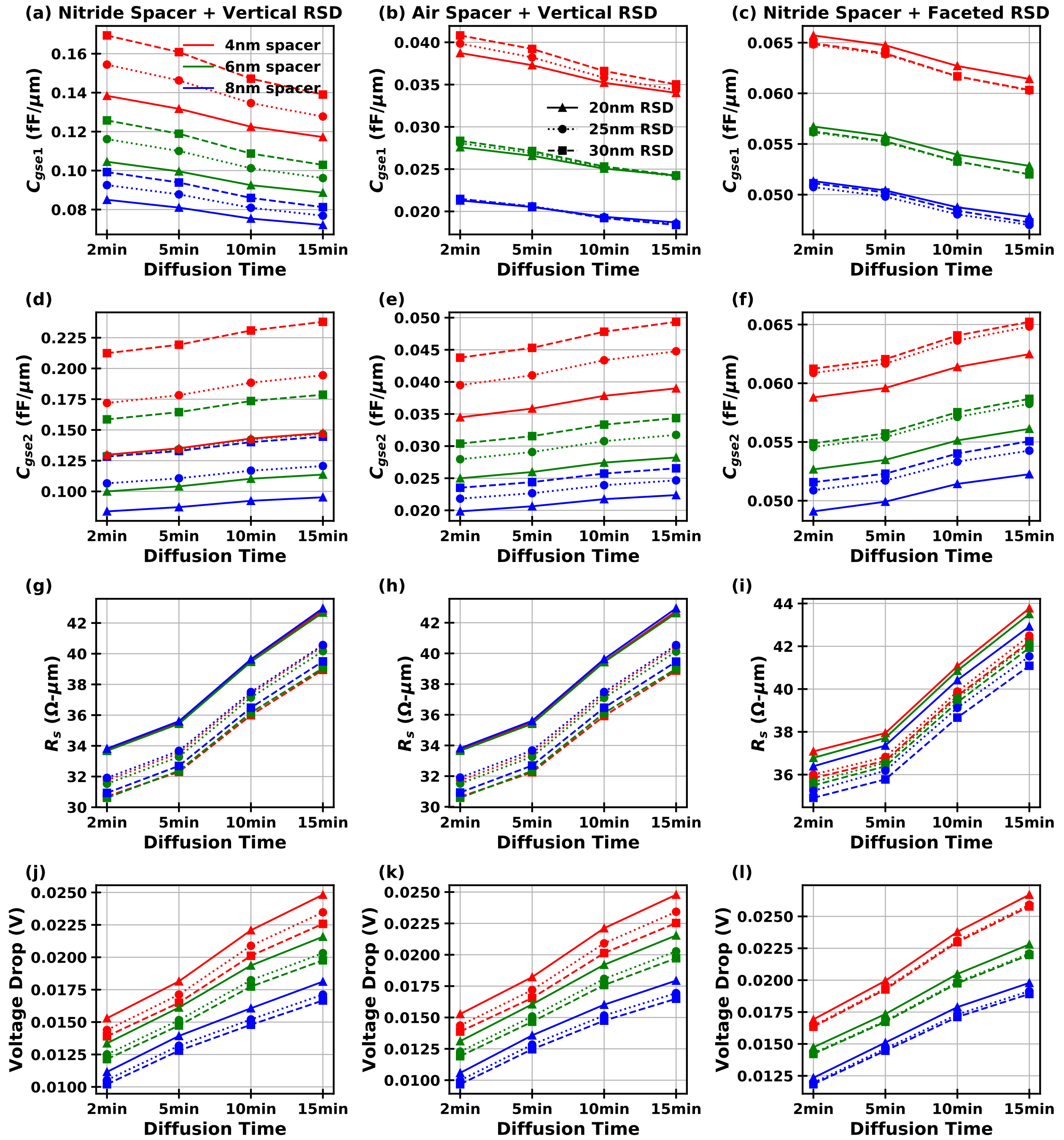


Figure 5.13: Extract (a)-(c) C_{gse1}, (d)-(f) C_{gse2}, (g)-(i) r_s, and (j)-(l) the voltage drop on the RSD for the source side at V_{GSi}=0.45V and V_{DSi}=0.7V. The same column of subplots represents the same type of device structure as indicated in (a)-(c).

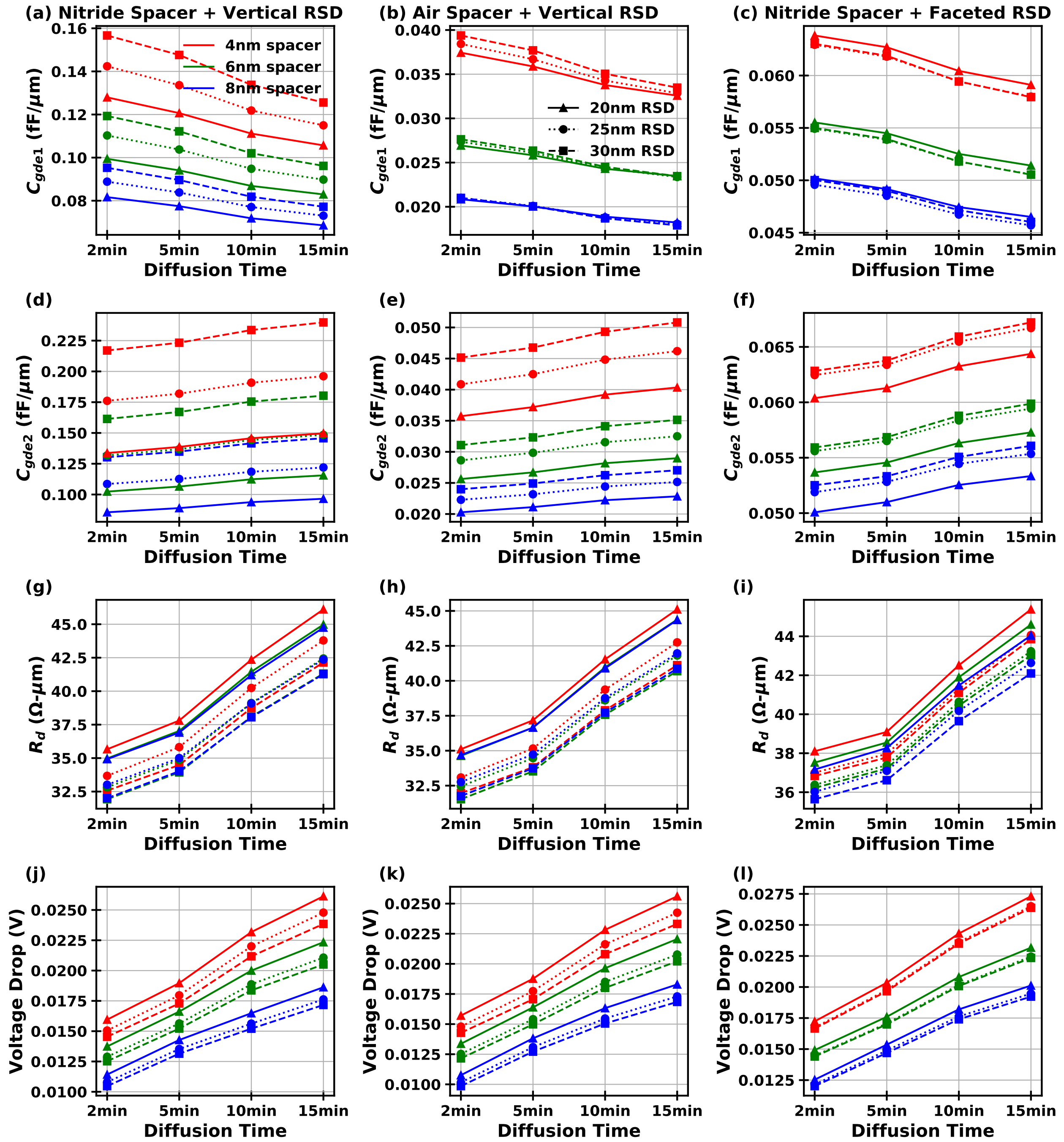


Figure 5.14: Extract (a)-(c) C_{gde1}, (d)-(f) C_{gde2}, and (g)-(i) r_d, and (j)-(l) the voltage drop on the RSD for the drain side at V_{GSi}=0.45V and V_{DSi}=0.7V. The same column of subplots represents the same type of device structure as indicated in (a)-(c).

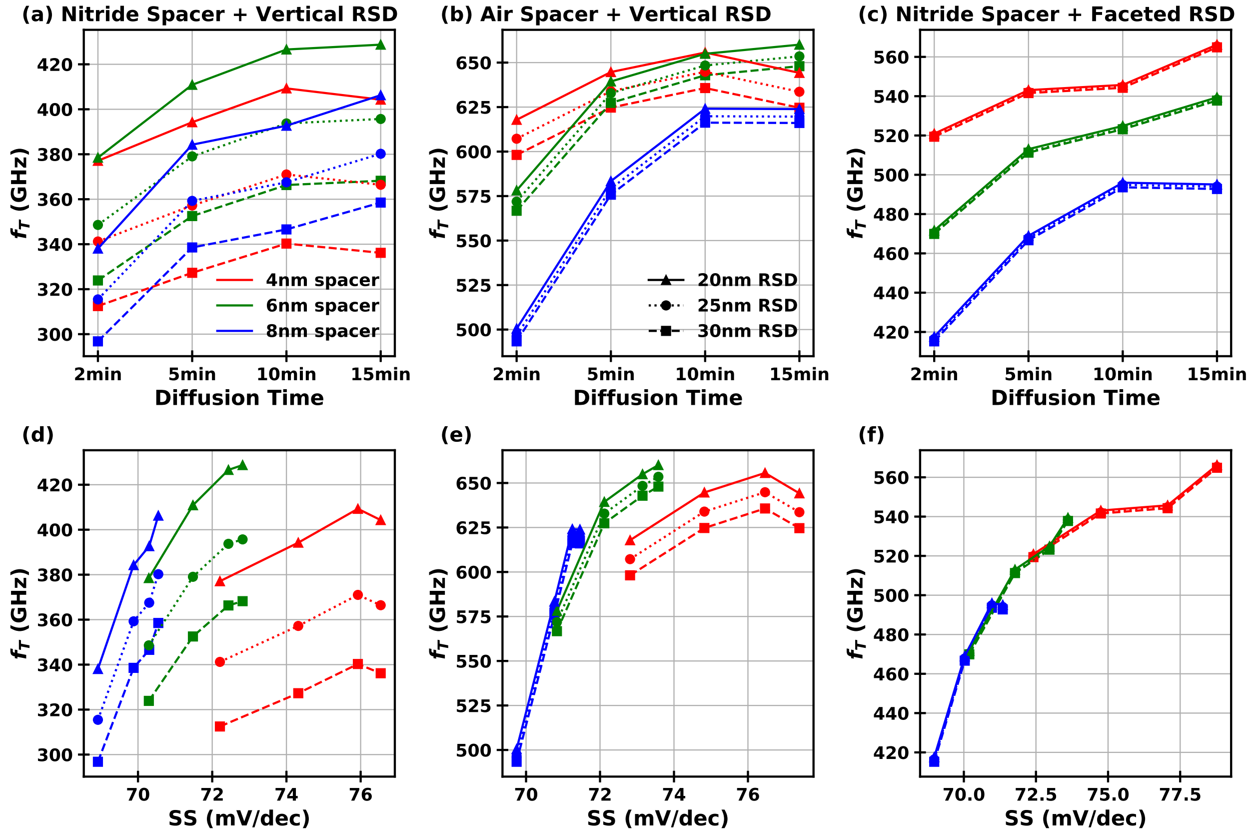


Figure 5.15: Extracted f_t at at V_{GSi}=0.45V and V_{DSi}=0.7V for various device structures and process parameters. Legends are shared by all the subplots. The same column of subplots represents the same type of device structure as indicated in (a)-(c).

(g)-(i).) Since a generally low series resistance can be achieved for the RSD structure, the performance degradation due to RSD is dominated by the parasitic capacitance. Therefore, a relatively thin RSD, a moderate spacer thickness, and a sufficient dopant diffusion time are preferred to reduce the parasitic capacitance while maintaining a high intrinsic performance. However, the diffusion time also affects the short channel effect (Fig. 5.11(g)-(i)) and the outer fringing field capacitance for the spacer extension regions (Fig. 5.12). It should be carefully controlled to ensure well-engineered spacer regions without causing significant S/D overlaps.

Fig. 5.15 summarizes the overall performance. For simulations with vertical RSD, the 6nm spacer thickness is ideal for maximizing the f_t because of the trade-offs between parasitic capacitance and g_{mi} (Fig. 5.15(a)(b),) while a narrower spacer is preferred for the faceted RSD structure because the thickness has a smaller impact on the RSD parasitic components (Fig. 5.15(c).) There also exists a trade-off between the SS and the g_{mi} and hence the f_{ti} through adjusting the spacer thickness and dopant diffusion time (Fig. 5.15(d)-(f).) For

vertical RSD structures, an increased spacer thickness helps reduce the SS given the same f_t. The simulations with 6nm air-gap spacers, 10∼15min dopant diffusion time, and 20nm-thick RSD achieves the highest f_t without sacrificing the SS (Fig. 5.15(b)(e)) within the explored design space. The design provides f_t over 650GHz, while the SS is about 73mV/dec. The simulations with conventional RSD and nitride spacer reaches f_t over 400GHz, which is about 20% improvements from the state of the art with a similar device structure at room temperature [86].

5.5 Conclusion

Process simulations combined with experimentally-calibrated device models are used to predict NCFET performance fabricated with modern nanoscale CMOS technology. The channel characteristics are realistically estimated through MC simulations which captures the quasi-ballistic carrier transport. The high intrinsic performance and the reduced susceptibility to parasitic capacitance due to the NC gate stack enable high f_t that outperforms the state of the art. Further design space explorations with different RSD structures indicates that f_t as high as 650GHz can be achieved in a structure optimized with air-gap spacers, which is suitable for the next-generation high-speed electronics.

Chapter 6

Dynamical Effect of Excess Carriers On SOI FeFET Memory Operations

6.1 Introduction

Prior works on FeFET devices have focused on modeling the physics of the ferroelectric layer to enable larger memory windows and thereby demonstrate concepts for high-performance memory alternatives [96, 97, 98, 99]. These works have emphasized the precise modeling of ferroelectric switching and domain dynamics through matching experimental parameters to fit either Preisach's or Landau-Khalatnikov's (LK) model [100]. As a result of this focus, the charges in the silicon channel are usually assumed to instantly redistribute and achieve the steady state. However, such assumptions about the channel dynamics may not hold for all devices, such as those exhibiting SOI characteristics. For example, in SOI devices, applying a negative V_G leads to large hole accumulation and while these carriers remain, the device experiences associated decay transients. This is the same phenomenon commonly observed in capacitor-less (0-C) DRAM cells [101, 102, 103]. Our work focuses on this aspect and hence on the interplay between the ferroelectric-induced electrostatics and the channel carrier dynamics in SOI-based FeFETs, a class of compact FeFETs that can potentially be integrated into advanced CMOS platforms. The physics-based simulations are targeted at understanding the channel physics which does not require precise modeling of the FE switching, and a single-domain LK-based model for the FE layer suits our purpose to generate a clean physical picture. Nevertheless, the insights drawn from our simulations are generally applicable to FE models with more complicated domain structures and dimensions [104, 105]. [1]

Our findings suggest that the GIDL-induced hole generation increases the electric field experienced by the ferroelectric layer, thereby accelerating the device's program (PGM) operation. The holes generated are also retained during the zero-bias hold and readout phases due to weak diffusion/recombination effects, causing reduced depolarization via car-

[1] The content of this chapter is based on [106].

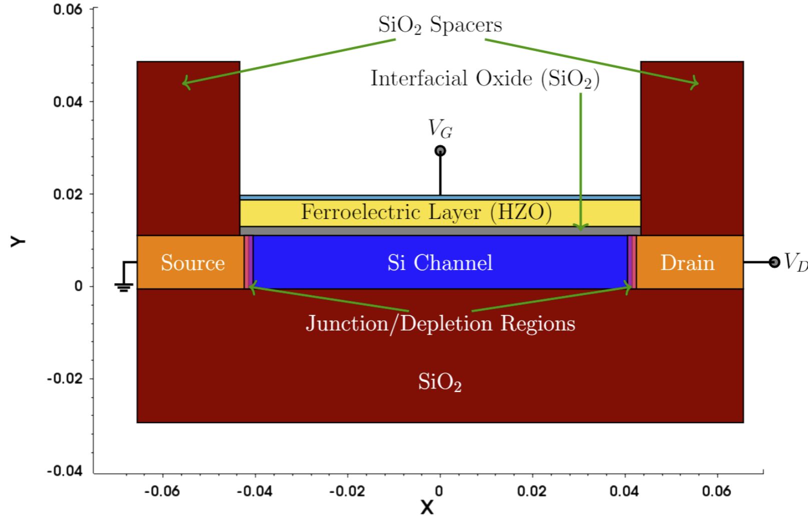


Figure 6.1: 2D schematic of the TCAD-simulated SOI-FeFET device.

rier screening [107, 108, 109]. The study of these excess hole dynamics and FE polarization transients provides new insight into the device physics governing the behavior of the FeFET, particularly regarding potential benefits utilizing the SOI structure. This in turn informs the future design of optimal FeFETs and their integration into modern CMOS technology as a superior memory alternative.

This chapter is organized as follows. In section 6.2, we discuss the simulation model and outline the FeFET device parameters. In section 6.3, we present our simulation results on the hole generation and retention, followed by further investigation on the device readout. We then discuss design takeaways drawn from our results in section 6.4. Finally, section 6.5 concludes the article.

6.2 FeFET Design and Simulation Model

We use Sentaurus TCAD [25] to simulate an n-type FDSOI FeFET shown in Fig. 6.1. The ferroelectric switching is modeled with the LK equation with a small domain-interaction term. The use of the LK model for domain switching in ultra-scaled ferroelectric devices was previously demonstrated [105]. The coefficients in the LK equation are set to match the parameters specified in table 6.1, which were selected based on experimental demonstrations of HfO_2 -based ferroelectrics reported previously [65, 67, 110, 111, 112].

Experimentally, material parameters such as the remanent polarization and permittivity can be tuned by varying the ferroelectric doping and deposition methods [110, 113, 114]. The ferroelectric thickness and gate length are selected to follow our previous experimental results, and the gate metal work function is matched to tungsten (W). For the silicon channel, drift-diffusion carrier transport models are solved self-consistently with SRH recombination

Parameter	Value
Gate Length (L_G)	90 nm
Ferroelectric Thickness (t_{FE})	5 nm
Ferroelectric Permittivity (ε_{FE})	27.3 ε_0
Coercive Field (E_C)	$1\,\mathrm{MV\,cm^{-1}}$
Remanant Polarization (P_r)	$4\,\mathrm{\mu C\,cm^{-2}}$
Channel Doping (N_{CH})	1×10^{17} (p)
Source/Drain Doping (N_S, N_D)	1×10^{20} (n)
Interfacial Oxide Thickness (t_{ox})	0.7 nm
SOI Thickness (t_{Si})	12 nm
Contact Resistance (R)	$1\,\mathrm{k\Omega - \mu m}$

Table 6.1: Selected device paramters

and non-local path band-to-band tunneling. The SRH recombination lifetime in our setting is 2 ns for both electrons and holes to give a conservative estimate of excess carrier lifetime.

6.3 Results and Observations

To demonstrate low-power operations, we use relatively small PGM and erase (ERS) voltages of ±1.5V in our simulations. Here, we define our PGM state as the low-current, high-V_T "OFF" state when a negative V_{GS} is applied, and vice-versa for the ERS state. The device is first programmed for 5 µs to ensure saturated polarization. The required PGM/ERS time for this device varies with applied V_{GS} and V_{DS}, but is generally less than 3 µs. The device is then held at $V_{GS} = V_{DS} = 0\,\mathrm{V}$ (defined as the "zero-bias hold" or interchangeably as the "hold" stage). The time for which this bias is maintained will be henceforth referred to as the "zero-bias-hold time" or interchangeably as the "hold time" (T_{H0}).

GIDL-Induced Hole Accumulation

Fig. 6.2 shows the I_D-V_G characteristics of our simulated device after different hold times. The waveform used here follows the fast V_G-ramp readout scheme described in Ref.[115]. A notable observation from Fig. 6.2 is the shift in V_T with hold time during the high-V_T state, while the threshold voltage during the low-V_T state remains constant for hold times sufficient for the polarization to switch in our settings (in this case, roughly greater than 1 µs). The shift in V_T suggests unusual carrier dynamics, which prompts us to look into the carrier distribution and electrostatics during the PGM phase.

Fig. 6.3a shows the lateral band diagram after PGM at a region close to the silicon/oxide interface. Due to the large electric field at the interface induced by the −1.5 V gate bias and the 0.5 V drain bias, a band-to-band (B2B) tunneling path is generated at the channel-

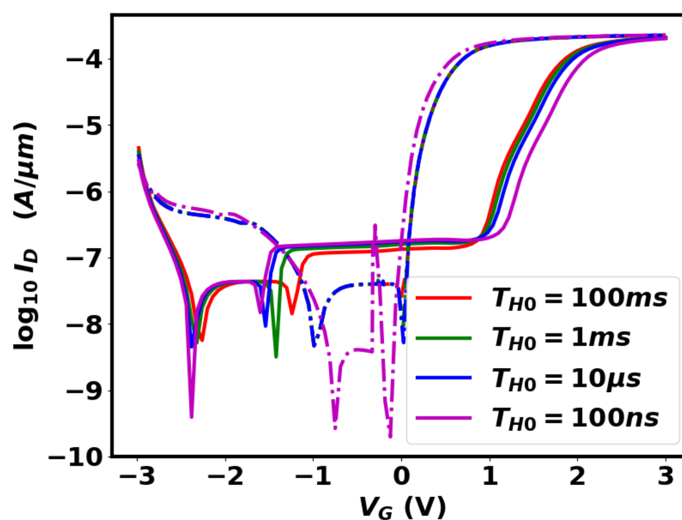
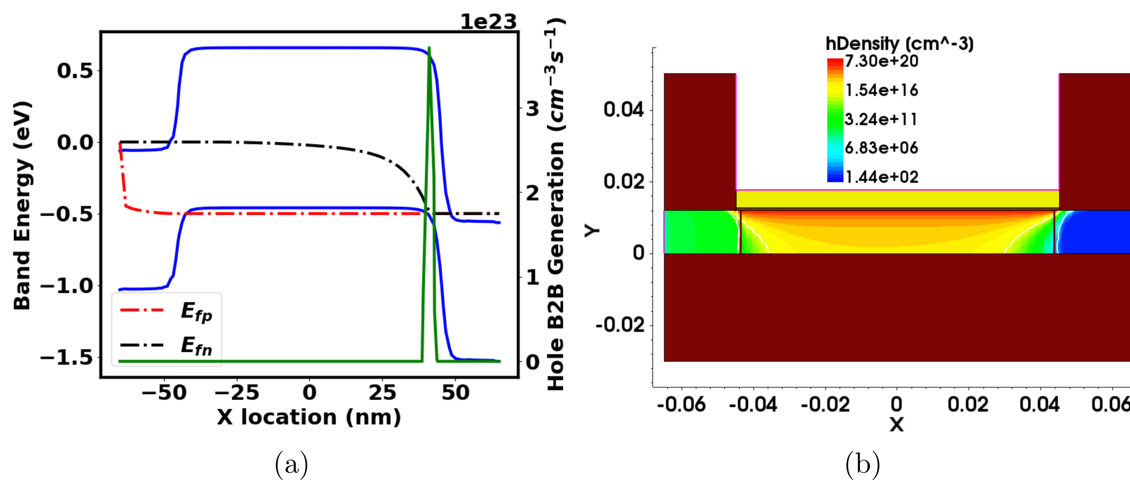


Figure 6.2: I_D-V_G readout after different zero-bias-hold times. Solid: high-V_T phase. Dashed: low-V_T phase. The rapid waveform ramp time of $5\,\mathrm{ns\,V^{-1}}$ ensures no polarization switching occurs during the readout. The magenta curve shows slightly different polarization values than the other curves due to a short hold time which is insufficient for the ferroelectric to reach its equilibrium polarization at $V_{GS} = 0\,\mathrm{V}$.



(a) (b)

Figure 6.3: (a) Band energies (blue) and hole band-to-band generation rate (green) of a horizontal cross-section taken in the silicon channel (Y= 11 nm, 1 nm from the interfacial oxide.) A large hole generation is observed in the channel close to drain, due to band-to-band tunneling. (b) A 2D plot showing the hole concentration of the entire device. A large hole concentration on the order of $10^{20}\,\mathrm{cm^{-3}}$ is observed at the cross section where (a) is taken.

drain edge [116]. Since the channel hole quasi-Fermi level (E_{fp}) is higher than the drain electron quasi-Fermi level (E_{fn}) at the bias condition, there is a tunneling of electrons from the channel valence band into the drain conduction band, or effectively, hole tunneling into the channel. As the channel hole concentration increases, channel valence band energy (E_v) and E_{fp} are brought down, reducing the junction field and the carrier distribution difference between the channel and the drain. The device eventually reaches a steady state where the hole injection through tunneling balances with the small hole outflow into the source through diffusion, as indicated by the proximity of the quasi-Fermi levels in Fig. 6.3a at the end of PGM. As a result of the SOI floating body effect, the channel hole concentration (Fig. 6.3b) reaches an extremely high value of $7.3 \times 10^{20}\,\text{cm}^{-3}$ at the channel/gate stack interface, far exceeding the near-equilibrium carrier statistics of a bulk device. This provides a boost to the programming electric field on the ferroelectric layer, resulting in a larger polarization window at a given PGM voltage. Notably, this GIDL-induced hole accumulation is also observed for low-V_{DS} PGM operations ($V_{DS} = 0.05\,\text{V}$). When the band-to-band tunneling model is turned off in the simulation, the electric field during the PGM phase becomes insufficient for the ferroelectric polarization to switch, which proves the importance of hole accumulation to enable low-voltage switching for SOI FeFET. The GIDL-induced hole accumulation in SOI devices has been observed experimentally, and is known to be the main operation principle for capacitor-less DRAM technology [102, 101]. Due to this large number of excess holes, the effect of hole dynamics on the FeFET device operation is not negligible for the PGM state.

Extended Hole Retention Time

To investigate the effect of hole dynamics on the device operation, we adapt a transient program-readout pulse sequence, shown in Fig. 6.4. A readout V_{GS} of 1 V ($V_{DS} = 0.5\,\text{V}$) is selected to give the maximum read margin based on the I_D-V_G characteristics in Fig. 6.2. The device is then held at $V_{GS} = V_{DS} = 0\,\text{V}$ before the readout pulse is applied. The voltage waveform described here is similar to that used in Section III A, but with a constant V_{GS} pulse for readout. The carrier dynamics during the zero-bias hold are examined in Fig. 6.5. During the hold in the low-V_T (i.e. ERS) state, the time required to reach steady-state hole concentration and FE polarization mainly varies with ρ, a term describing the polarization-switching time in the LK equation [117, 118, 57, 119]. This trend suggests that polarization switching is the dominant mechanism in the ERS state. In the high-V_T phase, however, the excess holes generated by GIDL during the PGM operation are retained throughout the 1 s hold (Fig. 6.5a). Unlike in the low-V_T phase, the high hole concentration is not strongly affected by ferroelectric switching, as seen from its invariance with ρ. The polarization boost for the PGM state due to the excess holes is retained during the hold, but slowly decays (Fig. 6.5b) as the hole concentration decays.

In contrast, no excess-carrier related evolution was observed for FE polarization in the ERS state which is associated with low hole concentrations. The large hole retention time can be explained by the weak channel hole leakage effects in the activated models. The hole

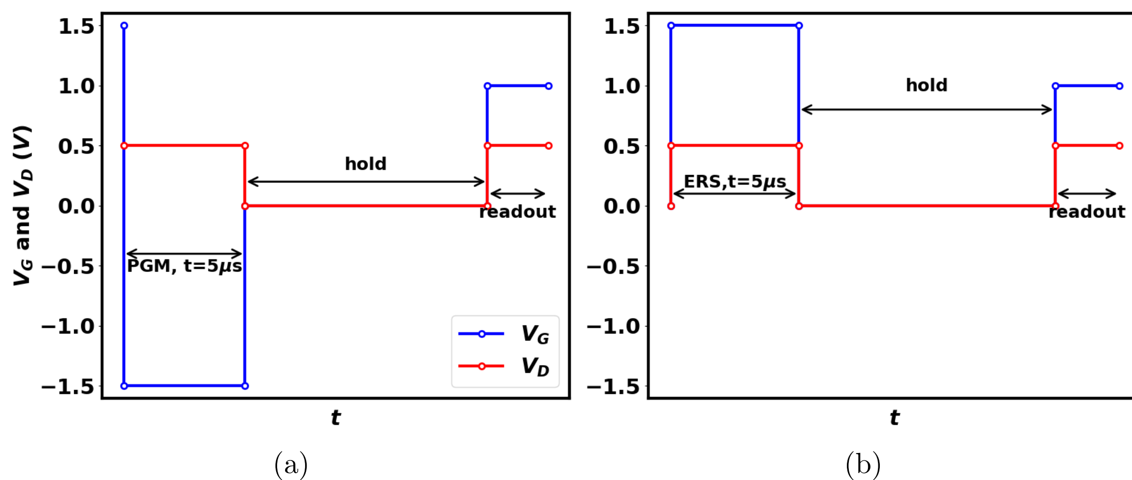


Figure 6.4: Waveform schematics for writing and reading (a) high-V_T (PGM) and (b) low-V_T (ERS) states. The hold operation is defined as the stage where $V_{GS} = 0\,V$. The readouts for both states are conducted at $V_{GS} = 1\,V$ and $V_{DS} = 0.5\,V$. The read time is on the order of ns.

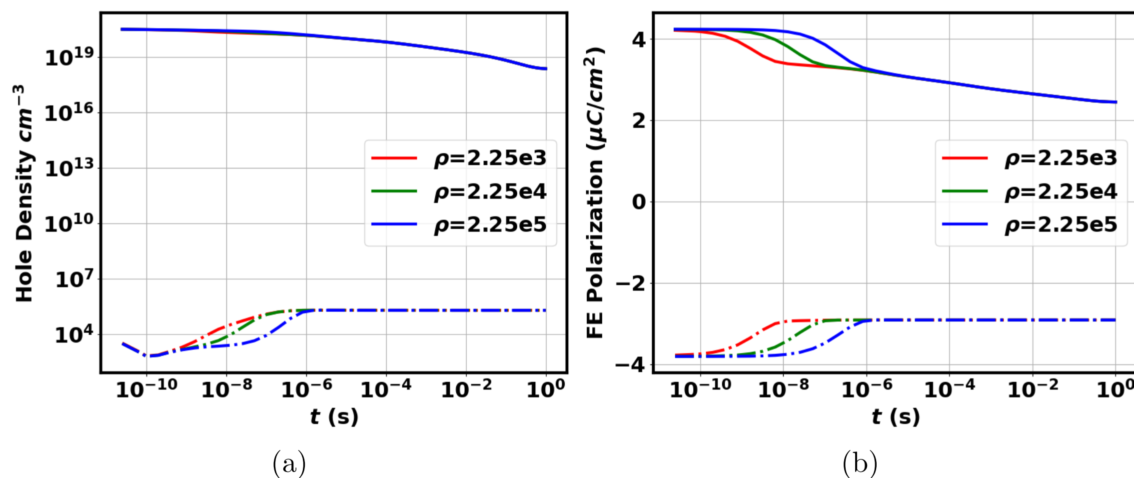


Figure 6.5: Transient evolution of (a) Hole density in the Si channel and (b) Ferroelectric polarization during zero-bias hold with different ρ. Solid curve: high-V_T phase, dashed curve: low-V_T phase. Reduced hole screening causes the polarization to degrade during hold of high-V_T phase, while the polarization during the low-V_T phase remains at a constant value during longer holds.

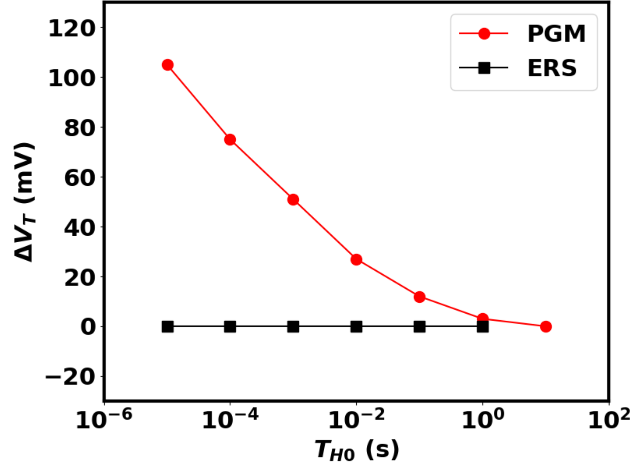


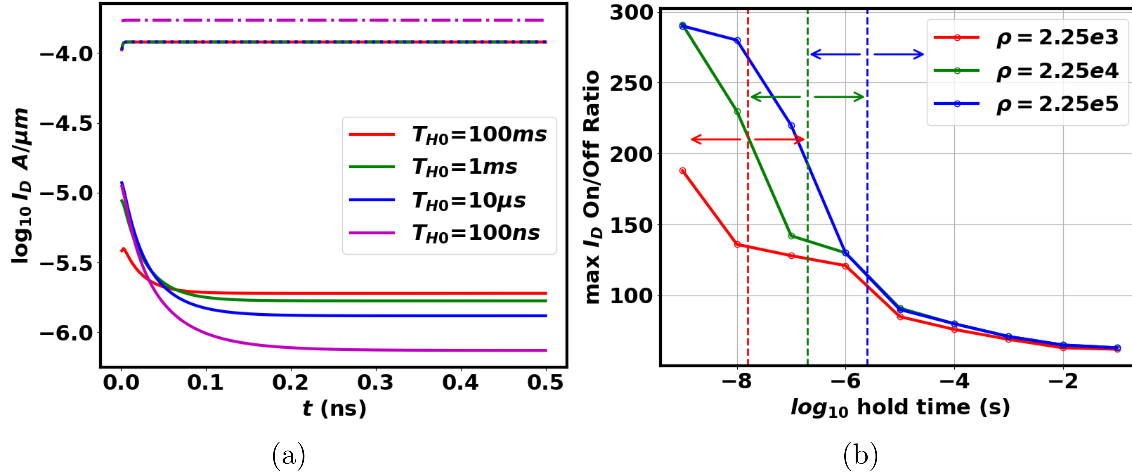
Figure 6.6: A demonstration of the device's V_T shift; as the zero-bias hold time T_{H0} increases from 10 µs to 1 s, the PGM V_T (red) steadily decreases from 1.12 V to 1.02 V, while the ERS V_T (black) remains unchanged.

diffusion current through the source and drain is limited by the potential barrier formed at the body p-n junction interface and hence remains low during the hold. The band-to-band recombination is weak during the hold due to the insufficient electric field at the channel-drain junction. Finally, even though the SRH recombination lifetime is set on the order of ns, we observe a hole retention time far beyond the nominal carrier lifetime as the carrier recombination is dominated by the low electron concentration. The recombination rate is further reduced by changes in barrier heights, which will be explained in the next subsection.

For PGM state during hold, the reduction of channel holes (Fig. 6.5a) should lead to increased conduction band energy at a given bias and hence increased $\mathrm{V_T}$, but the corresponding reduction of the FE polarization (Fig. 6.5b) is expected to have the opposite effect. The polarization effect on the $\mathrm{V_T}$ outweighs the hole electrostatic effects in our setting as the overall $\mathrm{V_T}$ decreases during hold (Fig. 6.6) for T_{H0} in a regime that is dominated by excess hole dynamics. For ERS state, the $\mathrm{V_T}$ is unchanged at the corresponding T_{H0} as the device is in steady state. As a result, the memory window reduces during hold, which can also be interpreted as the excess holes resulting in temporary boost to the memory window.

Effect of Carrier Dynamics on the Device Readout

To further investigate the effect of hole dynamics, we plot the device readout transients after different zero-bias hold times. Fig. 6.7a shows the resulting read margins. While the readout current remains roughly flat in the low-$\mathrm{V_T}$ phase, there is a sub-ns delay in the high-$\mathrm{V_T}$ readout before reaching a quasi-steady-state current which remains almost unchanged in ns time scales. This delay can be explained by the hole diffusion-limited mechanism, resulting from the large number of channel excess holes during the readout (Fig. 6.8a). To



(a) (b)

Figure 6.7: (a) Log drain current during readout (defined as the readout current) after different zero-bias-hold times, $\rho = 2.25 \times 10^5$. Solid: high-V_T , dashed: low-V_T . The low-V_T readout currents for hold times greater than the polarization switching time are identical and remain flat throughout. The plot is zoomed-in to sub-ns for better visualization of the decay. (b) Maximum read margin, defined as the I_D ratio of the quasi-steady-state drain current for the two states in (a), as a function of hold time. The arrows to the right show the time scale dominated by the LK equation, while the arrows to the left indicate the hole-dynamics-dominated zone, particularly during the high-V_T phase.

understand the different carrier contributions to the readout current, we visualize the source hole and electron currents separately. Fig. 6.8b shows a large hole current from channel to source that is consistent with the initial decay in the overall readout, suggesting non-negligible hole diffusion outflow due to a large concentration gradient and reduced junction field. Fig. 6.9 further illustrates the hole diffusion-initiated carrier dynamics. As holes flow from the channel into the source, the reduction in the net positive charges causes an increase in the conduction band energy (E_c). Because the potential barrier rises, fewer electrons are injected from the source into the channel. As a result, the source electron current (Fig. 6.10a) continues to decrease until a quasi-steady-state hole concentration is reached. The barrier modulation effect also extends the hole lifetime during hold and readout. Since the electron statistics in the channel are dominated by the source Fermi level, the increase in the channel E_c results in a lowered channel electron concentration, hence reducing the carrier recombination rate.

As the hold time increases, the quasi-steady-state readout current in the high-V_T phase increases, which reduces the maximum read margin. This is explained by the effect of ferroelectric polarization on the source electron injection current. Fig. 6.7b shows the maximum read margin as a function of hold time, and these variables are plotted for different ρ values to distinguish the effect of polarization switching from the carrier dynamics. We find these

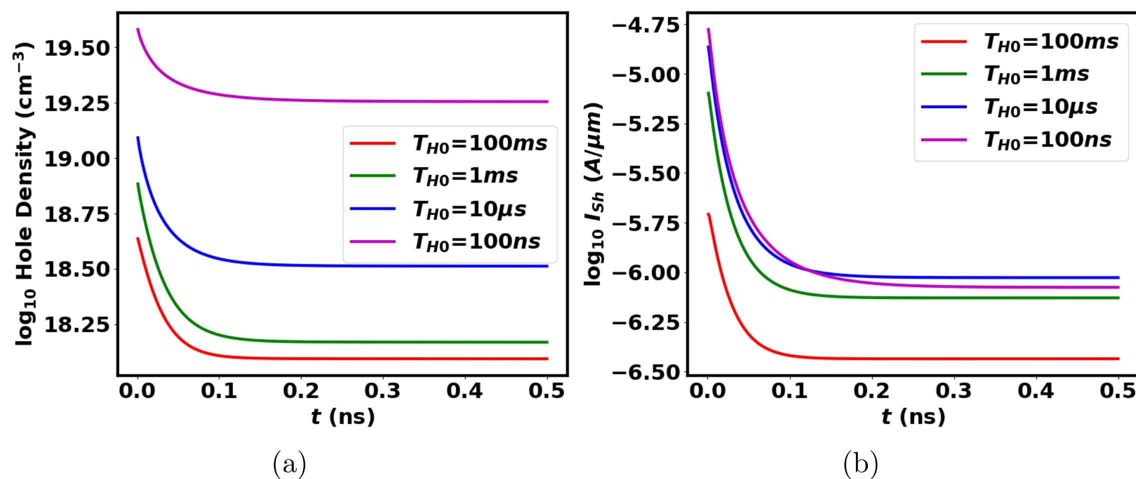


Figure 6.8: (a) Log hole density vs. time during readout in the high-V_T phase for different zero-bias-hold times. The hole density reaches quasi-steady state within 0.5 ns. (b) Hole source current during readout of the high-V_T phase. The hole current decreases with increasing hold time due to lower hole concentration, hence diffusion. The quasi-steady-state hole current at $T_{H0} = 100$ ns is slightly lower than $T_{H0} = 10$ µs because of incomplete polarization switching in the ferroelectric at short time scales.

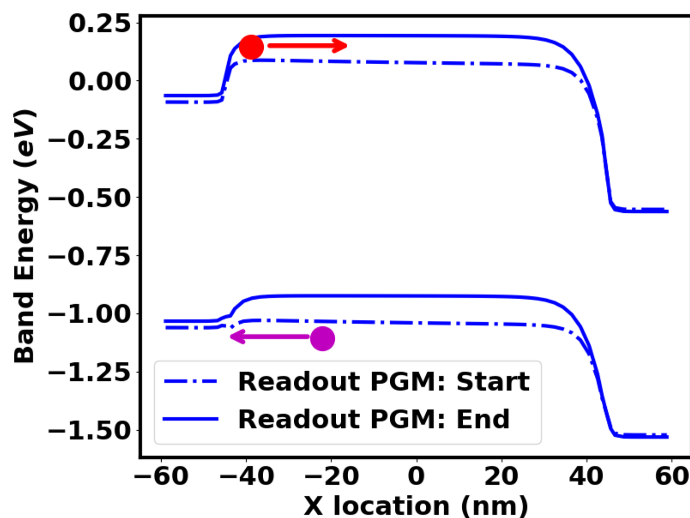


Figure 6.9: Energy band diagram at the start (dashed) and end (solid) of the high-V_T state readout. The increases as holes (purple) flow into the source, which increases the energy barrier for the injection of source electrons.

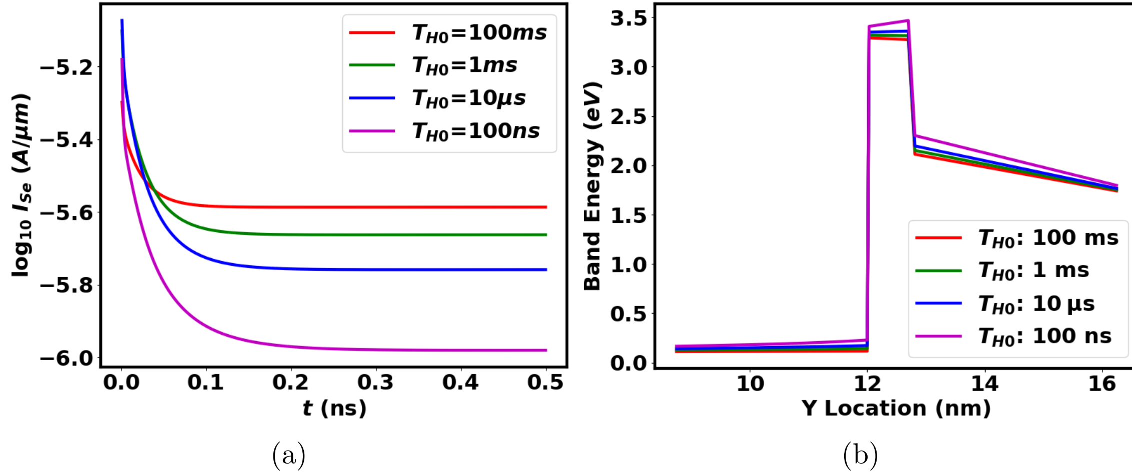


Figure 6.10: (a) Electron source current during readout of the high-V_T phase. The quasi-steady-state electron current increases with increasing zero-bias-hold time. (b) Conduction band energy of a vertical cross section showing the Si/SiO_2/FE stack at the end of high-V_T readout. The electron quasi-Fermi levels at the source are aligned for all the cases. The reduces as the hold time increases because of reduced polarization, which leads to an increased quasi-steady-state electron source current in (a).

characteristics highly correlated to the polarization difference between PGM and ERS states in Fig. 6.5b. The initial drops in the characteristics are due to the polarization responses to the change of V_{GS} from the write voltages to 0V, as governed by the LK model. For the PGM state, the polarization continues to decrease beyond the LK switching point because of reduced hole screening during the hold. This effect can be quite pronounced as seen in Fig. 6.5b in which the read margin is reduced by more than a half with the increase of hold time. The effect of polarization screening on the carrier transport during readout is further illustrated in Fig. 6.10b. As the hold time increases and the excess hold concentration reduces, the corresponding reduction in the ferroelectric polarization leads to a lowered conduction-band energy in the silicon channel and therefore a higher electron injection current from the source in quasi-steady state (Fig. 6.10a).

6.4 Design Perspectives

Benefits Due to the Hole Accumulation

The accumulation of holes in the channel enables low-V_{GS} switching, accelerates the PGM operation, and boosts the polarization window through the increased electric field experienced by the ferroelectric layer. Due to extended hole retention, the boosted polarization is retained throughout the hold stage. Despite the initial transient read delay due to excess

hole dynamics, the augmented polarization window offers some crucial design advantages over bulk devices with an MFIS gate stack, where the depolarization effect has been known to cause retention degradation due to incomplete charge screening [108]. In our SOI structure, the accumulated holes in the channel effectively compensate the depolarization field during the readout, leading to larger read margins over short hold times [107].

Strategies for Removing the Excess Carriers

Another important point of discussion is the ability to remove the excess holes in the channel during the high-V_T state, which may be desirable in order to avoid the read delay or to ensure a constant read margin regardless of hold time. One potential method is to create a tunneling path for hole outflow by applying a drain bias, which requires two conditions to hold. Firstly, the electric field at the drain edge has to be large enough, and typically band-to-band tunneling does not occur unless the channel E_v is higher than the drain E_c. Secondly, the channel E_{Fp} has to be lower than the drain E_{Fn} so that the net electron flow is from the drain conduction band to the channel valance band. A narrow window of V_{DS} can satisfy both conditions at the same time, as the first condition fails if V_{DS} is too low (Fig. 6.11a), and the second condition fails if V_{DS} is too high (Fig. 6.11c). Furthermore, a carefully chosen V_{DS} (Fig. 6.11b) will still result in a minimal hole outflow, and the outflow diminishes as the initial reduction of channel hole concentration results in raised E_{Fp}. In Fig. 6.11d, the hole concentration remains at the order of $10^{20} cm^{-3}$ during hold regardless of the V_{DS} . The small tunneling outflow at V_{DS} =0.8V contributes to negligible difference to the initial hole evolution, while the increased V_{DS} blocks the hole diffusion into the drain side and hence results in reduced total hole outflow in the long run. The results suggest that applying a large V_{DS} during hold will not be effective for removing holes in this device configuration, but might be a good way to retain memory window. Alternatively, the hole outflows through diffusion can be modulated through the gate voltage as demonstrated in III-C where an increased V_{GS} leads to rapid hole removal. Therefore, the excess holes removal during hold can be achieved through gate bias or work function engineering to reduce the potential barrier height for the holes in the PGM state.

Summary and Trade-Offs

In summary, the hole accumulation in SOI-FeFETs can give rise to potential benefits in the programming and retention of the device. To optimize these benefits in real designs, one needs to keep in mind the caveats regarding transient read delay and hold-time-dependent read margin caused by the excess carrier dynamics discussed in the previous section. By changing the bias conditions, it is also possible to control the carrier dynamics in other stages and hence extend the application space and efficiency of such devices. Nevertheless, effective control of the carrier dynamics will require more precise designs on the electrostatics and other physical parameters.

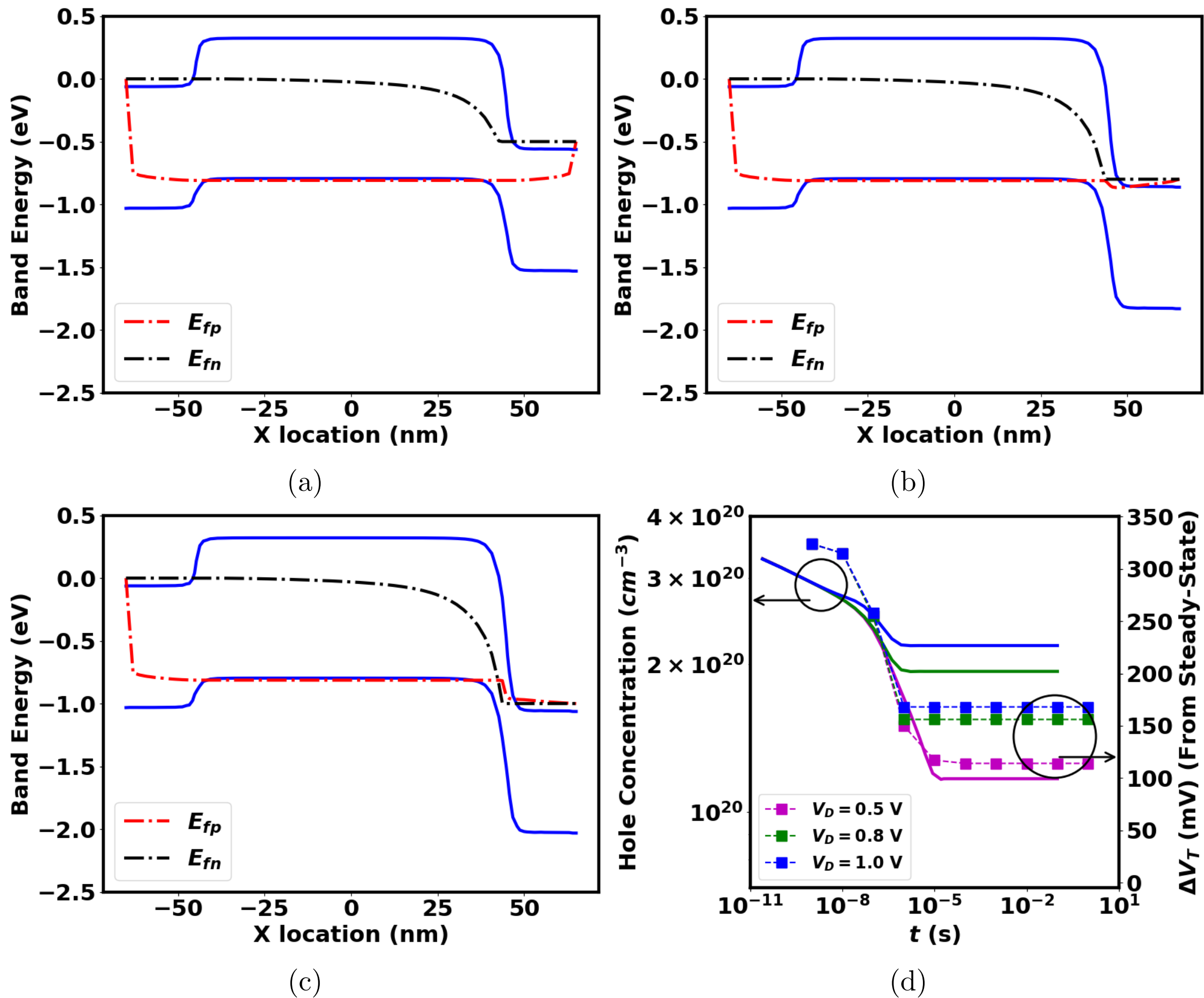


Figure 6.11: Lateral band diagrams with quasi-Fermi levels taken at Y= 0.011 during the 0 V_{GS} hold with applied drain bias. (a) $V_{DS} = 0.5\,V$ (b) $V_{DS} = 0.8\,V$, and (c) $V_{DS} = 1.0\,V$. (d) The channel hole concentration as a function of zero-bias hold time, for the 3 different V_D values. Note that the point at which the hole concentration begins to level off varies significantly with the chosen drain voltage. The high hole concentration and associated ΔV_t below 1 µs is related to incomplete polarization switching.

6.5 Conclusion

We have modeled an n-type SOI-FeFET in TCAD and observed GIDL-induced channel hole accumulation at a negative V_G for device programming. The GIDL-induced hole accumulation increases the ferroelectric electric field, and hence enables low-V_{GS} , fast PGM operations, and enhanced memory windows. The hole retention time is found beyond ms due to weak diffusion and recombination effects during the zero-bias hold. Hence, the extended hole retention, and thus the polarization boost, are carried over to the readout stage. The conclusions can be easily generalized to *p*-type SOI-FeFET memories which benefit from the extended retention of excess electrons. Despite some caveats on the initial readout delay and hold-dependent read margin, with careful considerations and optimizations on the electrostatics, designers can potentially utilize the carrier dynamics in this device configuration to obtain improved operation speed and memory retention.

Chapter 7

Conclusions

7.1 Chapters Summary

The main motivation of this dissertation is to understand the physics of the Ferroelectric and NC gate stack, to analyze its interplay with the channel electrostatics for scaled MOSFETs, and to shed lights on the device optimizations based on the understandings. Chapter 1 provides an overview of the background and prospects of the advanced gate stack researches. The fundamental principles of the negative capacitance effect and multi-domain considerations are discussed, followed by a review of the research progress for NCFETs and FeFETs with Hf-based ferroelectric gate stack. Such systems show promising prospects for both process and structural integration, enabling reduced-voltage operations.

Modeling of NCFET Experiments

In chapters 2, 3, and 4, combined modeling-experimental studies were focused on CMOS-compatible NCFET structures demonstrating potential amplification effects without hysteresis. In chapter 2, the fast response of the NC effect is confirmed through the Ring Oscillator (RO) modeling and experiments. The consistency between the device characteristics in DC and the RO transient response was demonstrated over multiple V_{DD} with per-stage delay down to less than 10ps. This is the experimental validation for the response speed of NCFETs at the shortest time scale that has been reported to date.

In chapters 3 and 4, we focused on anomalous transistor characteristics that are not explainable with classical "higher-κ" gate stack scaling trend. The experimental results described in chapter 3 are characterized by a C_g increase for the NCFET as the L_g reduces. This is in contrast with the C_g decrease due to subsurface leakages for FDSOI devices with linear gate dielectrics. The experimental results described in chapter 4 are characterized by a reversed trend for SS as a function of I_D, and the non-monotonic improvements of gate control from the controlled MOSFET to the NCFET as a function of L_g. TCAD models are developed to capture the polarization responses in the NC gate stacks through the calibration to controlled and NCFET experiments. For each of the studies, the TCAD

model successfully captured the difference between the controlled MOSFET and NCFET and quantitatively reproduces the experimental results over multiple L_g and V_{DS} through a Landau model with consistent parameters across all the NCFET simulations. The non-conventional trends can be explained with the non-linear potential amplification effect as described by the "S-shaped" polarization-electric field response of the FE layer, a distinct behaviour of the NC state that has been theoretically predicted and measured from simple material stack systems [21, 22].

NCFET Design

The non-linearity discovered in chapter 3 and 4 have important implications on the NCFET design. As the NC effect depends on the external electric field, an optimized NCFET should have the maximized potential amplification effect at the operation biases. We pointed out in chapter 4 that this could be achieved by introducing fixed charges or long-range dipoles into the gate stack, which effectively modulates the built-in electric field in the FE layer. When this can be achieved, the NC gate stack can lead to substantial improvements in ultra-scaled devices.

In chapter 5, the performance for short-channel SOI NCFET devices are projected with a realistic setting based on existing technologies, NCFET experimental results, and self-consistent device calculations with Monte-Carlo simulations for carrier transport. The improved EOT from the state of the art SOI technology and the non-degraded low-field mobility lead to f_t over 400GHz with the same type of spacer and raised source and drain modules as [86]. With the help of air-gap spacers, f_t over 650GHz is expected for the optimized NCFET. The relatively low susceptibility to the parasitic capacitance of the NCFET is the key enabler to the high-speed operation.

SOI FeFET Modeling and Design

In chapter 8, we pointed out the importance of the excess minority carrier in the program operation for SOI FeFET. The excess hole accumulation induced by the GIDL effect is a phenomenon specific to the n-type SOI MOSFET structure, and we investigated the opportunity to exploit it and enhance the memory writeability. We found that the effect can provide a significant boost to the electric field in the FE layer during programming, but is accompanied with hole transients that affect the read operation. Design insights are drawn from the findings, and strategies for device design with the awareness to the excess carrier effects are discussed. When the devices are designed properly, this unique property of SOI memory devices can be utilized to improve the FeFET write speed and retention.

7.2 Future Works

Extended Validation of NC Response Speed

The NC response speed study in chapter 2 is based on the distinguishable difference in the RO currents and delay that would otherwise not exist without the NC effect. The study in this dissertation was limited by the lower ON current of the RO NCFET devices than the control FinFET for large V_{GS}, making it unclear whether the NC effect persists at high V_{DD}. Therefore, the validation was focused on relatively low V_{DD} at which an NC effect could be confidently identified. To extend the validated delay and operation bias range, it would be ideal to test the RO that is made with well-engineered NCFETs that show clear NC effects at high biases in the DC measurements. This would allow for the validation to be confidently applied to high V_{DD} which corresponds to reduced per-stage delay. Alternatively, response speed of the NC effect can be studied through RF transistor measurements, and [84] demonstrated gain measurements with oscillations up to the order of 10GHz. Further optimizing the RF transistors with the design discussed in chapter 5 would reduce the effect of unwanted parasitics and hence enable probing of the NC speed at increased frequency.

Future Directions for NCFET Optimizations

The device optimizations in chapter 5 where focused on the performance in the ON state. The EOT was assumed to be a constant based on the CV measurements that could be captured by linear gate stack simulations. However, as shown in chapter 4, it is possible that non-linearity is present in the sub-threshold regime, and careful analysis of the complete transfer characteristics is required to identify these effects. When the OFF current is at stake, one may engineer the fixed charge of long-range dipoles in the gate stack to maximize the capacitance matching effect in the sub-threshold regime. Another possible way to further optimize the device is through the gate stack. While the gate stack is fixed in chapter 5, additional design parameters such as the IL thickness might have a significant impact on the performance. For example, a moderately reduced IL thickness through SiO_2 scavanging may degrade the mobility but enhance the intrinsic capacitance, and the overall effects for short channel devices have yet to be understood. Besides, enhanced NC effect stronger than that in the present study is possible and may allow for further scaling and hence performance improvements. Extensive experimental works on the CV and mobility are required for such investigations, but the methodology developed in this study will be generally applicable with the measured characteristics.

Realistic FeFET Modeling

The SOI FeFET study is focused on the channel carrier dynamics and its effect on the electric field in the gate stack, while the FE model was intentionally idealized to allow for a clean analysis of the effects of interest. While the conclusion was expected to be qualitatively

applicable to realistic FE memory, the impact of the excess carrier effect on multi-domain FE polarization states has not been quantitatively understood. The modeling of multi-domain FE states in scaled transistor itself is a topic that requires extensive researches, and it is the key to study and engineer the variations of FeFET memories and thus establishing relevance of this technology.

Concluding Remarks

The modeling and simulation works are intended to provide guidelines to design and understand future experiments. Therefore, the majority of the works in this dissertation should be substantiated by experiments which further test or utilize the models. These can also result in possible extensions or corrections of the models as discussed above.

Bibliography

[1] Kelin J. Kuhn. "Considerations for Ultimate CMOS Scaling". In: *IEEE Transactions on Electron Devices* 59.7 (2012), pp. 1813–1828. DOI: 10.1109/TED.2012.2193129.

[2] *International Roadmap for Devices and Systems (IRDS^{TM}) 2021 Edition*. 2021. URL: https://irds.ieee.org/editions/2021.

[3] Dragica Vasileska et al. *Schred*. Feb. 2006. DOI: doi : 10 . 4231 / D34T6F441. URL: https://nanohub.org/resources/schred.

[4] Shahab Siddiqui et al. "Chapter 5 - Limits of Gate Dielectrics Scaling". In: *Handbook of Thin Film Deposition (Fourth Edition)*. Ed. by Krishna Seshan and Dominic Schepis. Fourth Edition. William Andrew Publishing, 2018, pp. 107–145. ISBN: 978-0-12-812311-9. DOI: https://doi.org/10.1016/B978-0-12-812311-9.00005-0. URL: https://www.sciencedirect.com/science/article/pii/B9780128123119000050.

[5] Takashi Ando. "Ultimate Scaling of High-κ Gate Dielectrics: Higher-κ or Interfacial Layer Scavenging?" In: *Materials* 5.3 (2012), pp. 478–500. ISSN: 1996-1944. DOI: 10.3390/ma5030478. URL: https://www.mdpi.com/1996-1944/5/3/478.

[6] T. Ando et al. "Understanding mobility mechanisms in extremely scaled HfO¡inf¿2¡/inf¿ (EOT 0.42 nm) using remote interfacial layer scavenging technique and V¡inf¿t¡/inf¿-tuning dipoles with gate-first process". In: *2009 IEEE International Electron Devices Meeting (IEDM)*. 2009, pp. 1–4. DOI: 10.1109/IEDM.2009.5424335.

[7] Takashi Ando et al. "Physical origins of mobility degradation in extremely scaled SiO2/HfO2 gate stacks with La and Al induced dipoles". In: *Applied Physics Letters* 96.13 (2010), p. 132904. DOI: 10.1063/1.3373914. eprint: https://doi.org/10.1063/1.3373914. URL: https://doi.org/10.1063/1.3373914.

[8] E. Cartier et al. "Fundamental aspects of HfO¡inf¿2¡/inf¿-based high-k metal gate stack reliability and implications on t¡inf¿inv¡/inf¿-scaling". In: *2011 International Electron Devices Meeting*. 2011, pp. 18.4.1–18.4.4. DOI: 10.1109/IEDM.2011.6131579.

[9] Sayeef Salahuddin and Supriyo Datta. "Use of Negative Capacitance to Provide Voltage Amplification for Low Power Nanoscale Devices". In: *Nano Letters* 8.2 (2008). PMID: 18052402, pp. 405–410. DOI: 10.1021/nl071804g. eprint: https://doi.org/10.1021/nl071804g. URL: https://doi.org/10.1021/nl071804g.

[10] A.F. Devonshire. “Theory of ferroelectrics”. In: *Advances in Physics* 3.10 (1954), pp. 85–130. DOI: 10.1080/00018735400101173. eprint: https://doi.org/10.1080/00018735400101173. URL: https://doi.org/10.1080/00018735400101173.

[11] L.D. LANDAU and E.M. LIFSHITZ. “CHAPTER II - ELECTROSTATICS OF DIELECTRICS”. In: *Electrodynamics of Continuous Media (Second Edition)*. Ed. by L.D. LANDAU and E.M. LIFSHITZ. Second Edition. Vol. 8. Course of Theoretical Physics. Amsterdam: Pergamon, 1984, pp. 34–85. ISBN: 978-0-08-030275-1. DOI: https://doi.org/10.1016/B978-0-08-030275-1.50008-4. URL: https://www.sciencedirect.com/science/article/pii/B9780080302751500084.

[12] J. C. Wong and S. Salahuddin. “Negative Capacitance Transistors”. In: *Proceedings of the IEEE* 107.1 (Jan. 2019), pp. 49–62. ISSN: 0018-9219. DOI: 10.1109/JPROC.2018.2884518.

[13] Michael Hoffmann et al. “On the stabilization of ferroelectric negative capacitance in nanoscale devices”. In: *Nanoscale* 10 (23 2018), pp. 10891–10899. DOI: 10.1039/C8NR02752H. URL: http://dx.doi.org/10.1039/C8NR02752H.

[14] A. M. Bratkovsky and A. P. Levanyuk. “Very large dielectric response of thin ferroelectric films with the dead layers”. In: *Phys. Rev. B* 63 (13 Mar. 2001), p. 132103. DOI: 10.1103/PhysRevB.63.132103. URL: https://link.aps.org/doi/10.1103/PhysRevB.63.132103.

[15] Pavlo Zubko et al. “Negative capacitance in multidomain ferroelectric superlattices”. In: *Nature* 534.7608 (June 2016), pp. 524–528. ISSN: 1476-4687. DOI: 10.1038/nature17659. URL: https://doi.org/10.1038/nature17659.

[16] Atanu K. Saha and Sumeet K. Gupta. “Multi-Domain Negative Capacitance Effects in Metal-Ferroelectric-Insulator-Semiconductor/Metal Stacks: A Phase-field Simulation Based Study”. In: *Scientific Reports* 10.1 (June 2020), p. 10207. ISSN: 2045-2322. DOI: 10.1038/s41598-020-66313-1. URL: https://doi.org/10.1038/s41598-020-66313-1.

[17] Ajay K. Yadav et al. “Spatially resolved steady-state negative capacitance”. In: *Nature* 565.7740 (Jan. 2019), pp. 468–471. ISSN: 1476-4687. DOI: 10.1038/s41586-018-0855-y. URL: https://doi.org/10.1038/s41586-018-0855-y.

[18] T. Rollo et al. “Stabilization of negative capacitance in ferroelectric capacitors with and without a metal interlayer”. In: *Nanoscale* 12 (10 2020), pp. 6121–6129. DOI: 10.1039/C9NR09470A. URL: http://dx.doi.org/10.1039/C9NR09470A.

[19] Asif I. Khan et al. “Differential voltage amplification from ferroelectric negative capacitance”. In: *Applied Physics Letters* 111.25 (2017), p. 253501. DOI: 10.1063/1.5006958. eprint: https://doi.org/10.1063/1.5006958. URL: https://doi.org/10.1063/1.5006958.

[20] Asif Islam Khan et al. “Experimental evidence of ferroelectric negative capacitance in nanoscale heterostructures”. In: *Applied Physics Letters* 99.11 (2011), p. 113501. DOI: 10.1063/1.3634072. eprint: https://doi.org/10.1063/1.3634072. URL: https://doi.org/10.1063/1.3634072.

[21] M. Hoffmann et al. “Demonstration of High-speed Hysteresis-free Negative Capacitance in Ferroelectric Hf¡inf¿0.5¡/inf¿Zr¡inf¿0.5¡/inf¿O¡inf¿2¡/inf¿”. In: *2018 IEEE International Electron Devices Meeting (IEDM)*. 2018, pp. 31.6.1–31.6.4. DOI: 10.1109/IEDM.2018.8614677.

[22] Michael Hoffmann et al. “Unveiling the double-well energy landscape in a ferroelectric layer”. In: *Nature* 565.7740 (Jan. 2019), pp. 464–467. ISSN: 1476-4687. DOI: 10.1038/s41586-018-0854-z. URL: https://doi.org/10.1038/s41586-018-0854-z.

[23] David Esseni and Riccardo Fontanini. “Macroscopic and microscopic picture of negative capacitance operation in ferroelectric capacitors”. In: *Nanoscale* 13 (21 2021), pp. 9641–9650. DOI: 10.1039/D0NR06886A. URL: http://dx.doi.org/10.1039/D0NR06886A.

[24] A. Cano and D. Jiménez. “Multidomain ferroelectricity as a limiting factor for voltage amplification in ferroelectric field-effect transistors”. In: *Applied Physics Letters* 97.13 (2010), p. 133509. DOI: 10.1063/1.3494533. eprint: https://doi.org/10.1063/1.3494533. URL: https://doi.org/10.1063/1.3494533.

[25] *Sentaurus^{TM} Device User Guide, Version O-2018.06*. Synopsys, Inc, June 2018.

[26] T. S. Böscke et al. “Ferroelectricity in hafnium oxide thin films”. In: *Applied Physics Letters* 99.10 (2011), p. 102903. DOI: 10.1063/1.3634052. eprint: https://doi.org/10.1063/1.3634052. URL: https://doi.org/10.1063/1.3634052.

[27] J. Müller et al. “Ferroelectric hafnium oxide: A CMOS-compatible and highly scalable approach to future ferroelectric memories”. In: *2013 IEEE International Electron Devices Meeting*. 2013, pp. 10.8.1–10.8.4. DOI: 10.1109/IEDM.2013.6724605.

[28] Z. Krivokapic et al. “14nm Ferroelectric FinFET technology with steep subthreshold slope for ultra low power applications”. In: *2017 IEEE International Electron Devices Meeting (IEDM)*. Dec. 2017, pp. 15.1.1–15.1.4. DOI: 10.1109/IEDM.2017.8268393.

[29] Suraj S. Cheema et al. “Enhanced ferroelectricity in ultrathin films grown directly on silicon”. In: *Nature* 580.7804 (Apr. 2020), pp. 478–482. ISSN: 1476-4687. DOI: 10.1038/s41586-020-2208-x. URL: https://doi.org/10.1038/s41586-020-2208-x.

[30] J. Zhou et al. “Ferroelectric HfZrOx Ge and GeSn PMOSFETs with Sub-60 mV/decade subthreshold swing, negligible hysteresis, and improved Ids”. In: *2016 IEEE International Electron Devices Meeting (IEDM)*. 2016, pp. 12.2.1–12.2.4.

[31] W. Chung, M. Si, and P. D. Ye. “Hysteresis-free negative capacitance germanium CMOS FinFETs with Bi-directional Sub-60 mV/dec”. In: *2017 IEEE International Electron Devices Meeting (IEDM)*. 2017, pp. 15.3.1–15.3.4.

[32] Z. Yu et al. “Negative capacitance 2D MoS2 transistors with sub-60mV/dec sub-threshold swing over 6 orders, 250 μA/μm current density, and nearly-hysteresis-free”. In: *2017 IEEE International Electron Devices Meeting (IEDM)*. 2017, pp. 23.6.1–23.6.4.

[33] D. Kwon et al. “Improved Subthreshold Swing and Short Channel Effect in FDSOI n-Channel Negative Capacitance Field Effect Transistors”. In: *IEEE Electron Device Letters* 39.2 (Feb. 2018), pp. 300–303. ISSN: 0741-3106. DOI: 10.1109/LED.2017.2787063.

[34] D. Kwon et al. “Negative Capacitance FET With 1.8-nm-Thick Zr-Doped HfO2 Oxide”. In: *IEEE Electron Device Letters* 40.6 (2019), pp. 993–996.

[35] D. Kwon et al. “Near Threshold Capacitance Matching in a Negative Capacitance FET With 1 nm Effective Oxide Thickness Gate Stack”. In: *IEEE Electron Device Letters* 41.1 (2020), pp. 179–182.

[36] Xiahan Sang et al. “On the structural origins of ferroelectricity in HfO2 thin films”. In: *Applied Physics Letters* 106.16 (2015), p. 162905. DOI: 10.1063/1.4919135. eprint: https://doi.org/10.1063/1.4919135. URL: https://doi.org/10.1063/1.4919135.

[37] Min Hyuk Park et al. “Ferroelectricity and Antiferroelectricity of Doped Thin HfO2-Based Films”. In: *Advanced Materials* 27.11 (2015), pp. 1811–1831. DOI: 10.1002/adma.201404531.

[38] Lun Xu et al. “Kinetic pathway of the ferroelectric phase formation in doped HfO2 films”. In: *Journal of Applied Physics* 122.12 (2017), p. 124104. DOI: 10.1063/1.5003918. eprint: https://doi.org/10.1063/1.5003918. URL: https://doi.org/10.1063/1.5003918.

[39] H. Ota et al. “Fully coupled 3-D device simulation of negative capacitance FinFETs for sub 10 nm integration”. In: *2016 IEEE International Electron Devices Meeting (IEDM)*. 2016, pp. 12.4.1–12.4.4.

[40] J. Seo, J. Lee, and M. Shin. “Analysis of Drain-Induced Barrier Rising in Short-Channel Negative-Capacitance FETs and Its Applications”. In: *IEEE Transactions on Electron Devices* 64.4 (2017), pp. 1793–1798.

[41] A. K. Saha et al. “Ferroelectric transistor model based on self-consistent solution of 2D Poisson’s, non-equilibrium Green’s function and multi-domain Landau Khalatnikov equations”. In: *2017 IEEE International Electron Devices Meeting (IEDM)*. 2017, pp. 13.5.1–13.5.4.

[42] Z. Dong and J. Guo. “A Simple Model of Negative Capacitance FET With Electrostatic Short Channel Effects”. In: *IEEE Transactions on Electron Devices* 64.7 (July 2017), pp. 2927–2934. ISSN: 0018-9383. DOI: 10.1109/TED.2017.2706182.

[43] G. Pahwa, A. Agarwal, and Y. S. Chauhan. "Numerical Investigation of Short-Channel Effects in Negative Capacitance MFIS and MFMIS Transistors: Subthreshold Behavior". In: *IEEE Transactions on Electron Devices* 65.11 (2018), pp. 5130–5136.

[44] J. P. Duarte et al. "Negative-Capacitance FinFETs: Numerical Simulation, Compact Modeling and Circuit Evaluation". In: *2018 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*. Sept. 2018, pp. 123–128. DOI: 10.1109/SISPAD.2018.8551641.

[45] J. Zhou et al. "Negative Differential Resistance in Negative Capacitance FETs". In: *IEEE Electron Device Letters* 39.4 (2018), pp. 622–625.

[46] H. Zhou et al. "Negative Capacitance, n-Channel, Si FinFETs: Bi-directional Sub-60 mV/dec, Negative DIBL, Negative Differential Resistance and Improved Short Channel Effect". In: *2018 IEEE Symposium on VLSI Technology*. 2018, pp. 53–54.

[47] G. Pahwa, A. Agarwal, and Y. S. Chauhan. "Numerical Investigation of Short-Channel Effects in Negative Capacitance MFIS and MFMIS Transistors: Above-Threshold Behavior". In: *IEEE Transactions on Electron Devices* 66.3 (2019), pp. 1591–1598. DOI: 10.1109/TED.2019.2892186.

[48] P. Su and W. You. "Electrostatic Integrity in Negative-Capacitance FETs – A Subthreshold Modeling Approach". In: *2019 IEEE International Electron Devices Meeting (IEDM)*. 2019, pp. 7.3.1–7.3.4. DOI: 10.1109/IEDM19573.2019.8993444.

[49] Mengwei Si et al. "Steep-slope hysteresis-free negative capacitance MoS2 transistors". In: *Nature Nanotechnology* 13.1 (2018), pp. 24–28. DOI: 10.1038/s41565-017-0010-1. URL: https://doi.org/10.1038/s41565-017-0010-1.

[50] Asif I. Khan et al. "Ferroelectric negative capacitance MOSFET: Capacitance tuning amp; antiferroelectric operation". In: *2011 International Electron Devices Meeting*. 2011, pp. 11.3.1–11.3.4. DOI: 10.1109/IEDM.2011.6131532.

[51] Masaharu Kobayashi et al. "Experimental study on polarization-limited operation speed of negative capacitance FET with ferroelectric HfO¡inf¿2¡/inf¿". In: *2016 IEEE International Electron Devices Meeting (IEDM)*. 2016, pp. 12.3.1–12.3.4. DOI: 10.1109/IEDM.2016.7838402.

[52] Zhi Cheng Yuan et al. "Switching-Speed Limitations of Ferroelectric Negative-Capacitance FETs". In: *IEEE Transactions on Electron Devices* 63.10 (2016), pp. 4046–4052. DOI: 10.1109/TED.2016.2602209.

[53] Yu-Sen Jiang et al. "Operation bandwidth of negative capacitance characterized by the frequency response of capacitance magnification in ferroelectric/dielectric stacks". In: *J. Mater. Chem. C* 9 (4 2021), pp. 1401–1409. DOI: 10.1039/D0TC04025H. URL: http://dx.doi.org/10.1039/D0TC04025H.

[54] C. Gastaldi et al. “Transient Negative Capacitance of Silicon-doped HfO¡inf¿2¡/inf¿ in MFMIS and MFIS structures: experimental insights for hysteresis-free steep slope NC FETs”. In: *2019 IEEE International Electron Devices Meeting (IEDM)*. 2019, pp. 23.5.1–23.5.4. DOI: 10.1109/IEDM19573.2019.8993523.

[55] X. Lyu et al. “First Direct Measurement of Sub-Nanosecond Polarization Switching in Ferroelectric Hafnium Zirconium Oxide”. In: *2019 IEEE International Electron Devices Meeting (IEDM)*. 2019, pp. 15.2.1–15.2.4. DOI: 10.1109/IEDM19573.2019.8993509.

[56] Cristobal Alessandri et al. “Switching Dynamics of Ferroelectric Zr-Doped HfO¡sub¿2¡/sub¿”. In: *IEEE Electron Device Letters* 39.11 (2018), pp. 1780–1783. DOI: 10.1109/LED.2018.2872124.

[57] Korok Chatterjee, Alexander John Rosner, and Sayeef Salahuddin. “Intrinsic speed limit of negative capacitance transistors”. In: *IEEE Electron Device Letters* 38.9 (2017), pp. 1328–1330. DOI: 10.1109/LED.2017.2731343.

[58] Daewoong Kwon et al. “Response Speed of Negative Capacitance FinFETs”. In: *2018 IEEE Symposium on VLSI Technology*. 2018, pp. 49–50. DOI: 10.1109/VLSIT.2018.8510626.

[59] Yu-Hung Liao et al. “Anomalously Beneficial Gate-Length Scaling Trend of Negative Capacitance Transistors”. In: *IEEE Electron Device Letters* 40.11 (2019), pp. 1860–1863. DOI: 10.1109/LED.2019.2940715.

[60] Yu-Hung Liao et al. “Electric Field-Induced Permittivity Enhancement in Negative-Capacitance FET”. In: *IEEE Transactions on Electron Devices* 68.3 (2021), pp. 1346–1351. DOI: 10.1109/TED.2021.3049763.

[61] Xueqing Li et al. “Design of Nonvolatile SRAM with Ferroelectric FETs for Energy-Efficient Backup and Restore”. In: *IEEE Transactions on Electron Devices* 64.7 (2017), pp. 3037–3040. DOI: 10.1109/TED.2017.2707664.

[62] Sumitha George et al. “Nonvolatile memory design based on ferroelectric FETs”. In: *2016 53nd ACM/EDAC/IEEE Design Automation Conference (DAC)*. 2016, pp. 1–6. DOI: 10.1145/2897937.2898050.

[63] Ekaterina Yurchuk et al. “Origin of the endurance degradation in the novel HfO2-based 1T ferroelectric non-volatile memories”. In: *2014 IEEE International Reliability Physics Symposium*. 2014, 2E.5.1–2E.5.5. DOI: 10.1109/IRPS.2014.6860603.

[64] J. Muller et al. “High endurance strategies for hafnium oxide based ferroelectric field effect transistor”. In: *2016 16th Non-Volatile Memory Technology Symposium (NVMTS)*. 2016, pp. 1–7. DOI: 10.1109/NVMTS.2016.7781517.

[65] Korok Chatterjee et al. “Self-Aligned, Gate Last, FDSOI, Ferroelectric Gate Memory Device With 5.5-nm Hf0.8Zr0.2O2, High Endurance and Breakdown Recovery”. In: *IEEE Electron Device Letters* 38.10 (2017), pp. 1379–1382. DOI: 10.1109/LED.2017.2748992.

[66] Nanbo Gong and Tso-Ping Ma. "A Study of Endurance Issues in HfO2-Based Ferroelectric Field Effect Transistors: Charge Trapping and Trap Generation". In: *IEEE Electron Device Letters* 39.1 (2018), pp. 15–18. DOI: 10.1109/LED.2017.2776263.

[67] Ava Jiang Tan et al. "Ferroelectric HfO2 Memory Transistors with High-κ Interfacial Layer and Write Endurance Exceeding 10^{10} Cycles". In: *IEEE Electron Device Letters* (2021), pp. 1–1. DOI: 10.1109/LED.2021.3083219.

[68] M.H. Na et al. "The effective drive current in CMOS inverters". In: *Digest. International Electron Devices Meeting,* 2002, pp. 121–124. DOI: 10.1109/IEDM.2002.1175793.

[69] Y. Lin et al. "Analysis and Modeling of Inner Fringing Field Effect on Negative Capacitance FinFETs". In: *IEEE Transactions on Electron Devices* 66.4 (Apr. 2019), pp. 2023–2027. ISSN: 0018-9383. DOI: 10.1109/TED.2019.2899810.

[70] A. Rahman, S. Datta, and M. S. Lundstrom. "Theory of ballistic nanotransistors". In: *IEEE Transactions on Electron Devices* 50.9 (Sept. 2003), pp. 1853–1864. ISSN: 0018-9383. DOI: 10.1109/TED.2003.815366.

[71] Harshit Agarwal et al. "Proposal for capacitance matching in negative capacitance field-effect transistors". In: *IEEE Electron Device Letters* 40.3 (2019), pp. 463–466.

[72] A. Saeidi et al. "Double-Gate Negative-Capacitance MOSFET With PZT Gate-Stack on Ultra Thin Body SOI: An Experimentally Calibrated Simulation Study of Device Performance". In: *IEEE Transactions on Electron Devices* 63.12 (2016), pp. 4678–4684.

[73] Kamal Karda et al. "An anti-ferroelectric gated Landau transistor to achieve sub-60mV/dec switching at low voltage and high speed". In: *Applied Physics Letters* 106.16 (2015), p. 163501. DOI: 10.1063/1.4918649. eprint: https://doi.org/10.1063/1.4918649. URL: https://doi.org/10.1063/1.4918649.

[74] N. Pandey and Y. S. Chauhan. "Analytical Modeling of Short-Channel Effects in MFIS Negative-Capacitance FET Including Quantum Confinement Effects". In: *IEEE Transactions on Electron Devices* (2020), pp. 1–8.

[75] Junichi Hattori et al. "Fringing field effects in negative capacitance field-effect transistors with a ferroelectric gate insulator". In: *Japanese Journal of Applied Physics* 57.4S (Mar. 2018), 04FD07. DOI: 10.7567/jjap.57.04fd07.

[76] I Luk'yanchuk et al. "Harnessing ferroelectric domains for negative capacitance". In: *Communications Physics* 2.1 (2019), pp. 1–6.

[77] T. Cam et al. "Sustained Benefits of NCFETs Under Extreme Scaling to the End of the IRDS". In: *IEEE Transactions on Electron Devices* 67.9 (2020), pp. 3843–3851. DOI: 10.1109/TED.2020.3007398.

[78] Yasuhiro Abe et al. “Dipole formation at direct-contact HfO2/Si interface”. In: *Applied Physics Letters* 90.17 (2007), p. 172906. DOI: 10.1063/1.2731514. eprint: https://doi.org/10.1063/1.2731514. URL: https://doi.org/10.1063/1.2731514.

[79] Hiroaki Arimura et al. “Temperature-dependent La- and Al-induced dipole behavior monitored by femtosecond pump/probe photoelectron spectroscopy”. In: *Applied Physics Letters* 96.13 (2010), p. 132902. DOI: 10.1063/1.3374883. eprint: https://doi.org/10.1063/1.3374883. URL: https://doi.org/10.1063/1.3374883.

[80] Asif Islam Khan et al. “Work function engineering for performance improvement in leaky negative capacitance FETs”. In: *IEEE Electron Device Letters* 38.9 (2017), pp. 1335–1338.

[81] K. Toprasertpong et al. “ Asymmetric Polarization Response of Electrons and Holes in Si FeFETs: Demonstration of Absolute Polarization Hysteresis Loop and Inversion Hole Density over $2 \times 10^{13} cm^{-2}$ ”. In: *2020 IEEE Symposium on VLSI Technology*. June 2020, TF1.5.

[82] *International Roadmap for Devices and Systems (IRDS^{TM}) 2020 Edition*. 2020. URL: https://irds.ieee.org/editions/2020.

[83] Suraj Cheema, Nirmaan Shanker, and Li-Chen Wang et al. “Atomic-scale ferroic HfO2-ZrO2 superlattice gate stack for advanced transistors”. PREPRINT (Version 1) available at Research Square [https://doi.org/10.21203/rs.3.rs-413053/v1]. Apr. 2021.

[84] W. Li et al. “Demonstration of Low EOT Gate Stack and Record Transconductance on Lg=90 nm nFETs Using 1.8 nm Ferroic HfO_2-ZrO_2 Superlattice”. In: *2021 IEEE International Electron Devices Meeting (IEDM)*. 2021.

[85] K. Cheng et al. “Fully depleted extremely thin SOI technology fabricated by a novel integration scheme featuring implant-free, zero-silicon-loss, and faceted raised source/drain”. In: *2009 Symposium on VLSI Technology*. 2009, pp. 212–213.

[86] Jan Hoentschel et al. “22FDX® Technologies for Ultra-Low Power IoT, RF and mmWave Applications”. In: *Nanoelectronic Devices* 2.FDSOI (2019). ISSN: 2516-3914. DOI: 10.21494/ISTE.OP.2019.0321. URL: http://www.openscience.fr/22FDX-R-Technologies-for-Ultra-Low-Power-IoT-RF-and-mmWave-Applications.

[87] X. Li et al. “Selective Epitaxial Si:P Film for nMOSFET Application: High Phosphorous Concentration and High Tensile Strain”. In: *ECS Transactions* 64.6 (Aug. 2014), pp. 959–965. DOI: 10.1149/06406.0959ecst. URL: https://doi.org/10.1149/06406.0959ecst.

[88] K. Cheng et al. “Air spacer for 10nm FinFET CMOS and beyond”. In: *2016 IEEE International Electron Devices Meeting (IEDM)*. 2016, pp. 17.1.1–17.1.4. DOI: 10.1109/IEDM.2016.7838436.

[89] Yen-Kai Lin et al. "Spacer Engineering in Negative Capacitance FinFETs". In: *IEEE Electron Device Letters* 40.6 (2019), pp. 1009–1012. DOI: 10.1109/LED.2019.2911104.

[90] *Garand User Guide, Version S-2021.06*. Synopsys, Inc, June 2021.

[91] M. V. Fischetti and S. E. Laux. "Long-range Coulomb interactions in small Si devices. Part I: Performance and reliability". In: *Journal of Applied Physics* 89.2 (2001), pp. 1205–1231. DOI: 10.1063/1.1332423. eprint: https://doi.org/10.1063/1.1332423. URL: https://doi.org/10.1063/1.1332423.

[92] Tadayoshi Uechi, Takayuki Fukui, and Nobuyuki Sano. "3D Monte Carlo simulation including full Coulomb interaction under high electron concentration regimes". In: *physica status solidi (c)* 5 (Jan. 2008), pp. 102–106. DOI: 10.1002/pssc.200776547.

[93] Fareen Adeni Khaja. "Contact Resistance Improvement for Advanced Logic by Integration of Epi, Implant and Anneal Innovations". In: *MRS Advances* 4.48 (Oct. 2019), pp. 2559–2576. ISSN: 2059-8521. DOI: 10.1557/adv.2019.416. URL: https://doi.org/10.1557/adv.2019.416.

[94] Yang Liu et al. "A Tight-Binding Study of the Ballistic Injection Velocity for Ultrathin-Body SOI MOSFETs". In: *IEEE Transactions on Electron Devices* 55.3 (2008), pp. 866–871. DOI: 10.1109/TED.2007.915056.

[95] G. Ghibaudo. "New method for the extraction of MOSFET parameters". English. In: *Electronics Letters* 24 (9 Apr. 1988), 543–545(2). ISSN: 0013-5194. URL: https://digital-library.theiet.org/content/journals/10.1049/el_19880369.

[96] Chi-Yu Chan et al. "FeFET Memory Featuring Large Memory Window and Robust Endurance of Long-Pulse Cycling by Interface Engineering using High-k AlON". In: *2020 IEEE Symposium on VLSI Technology*. 2020, pp. 1–2. DOI: 10.1109/VLSITechnology18217.2020.9265103.

[97] K. Kaczmarek et al. "Understanding the memory window in 1T-FeFET memories: a depolarization field perspective". In: *2021 IEEE International Memory Workshop (IMW)*. 2021, pp. 1–4. DOI: 10.1109/IMW51353.2021.9439597.

[98] Chia-Sheng Hsu et al. "A Theoretical Study of Multidomain Ferroelectric Switching Dynamics With a Physics-Based SPICE Circuit Model for Phase-Field Simulations". In: *IEEE Transactions on Electron Devices* 67.7 (2020), pp. 2952–2959. DOI: 10.1109/TED.2020.2990891.

[99] Halid Mulaosmanovic et al. "Ferroelectric FETs With 20-nm-Thick HfO2 Layer for Large Memory Window and High Performance". In: *IEEE Transactions on Electron Devices* 66.9 (2019), pp. 3828–3833. DOI: 10.1109/TED.2019.2930749.

[100] Zhiyuan Fu et al. "Device Modeling and Application Simulation of Ferroelectric-FETS with Dynamic Multi-Domain Behavior". In: *2020 China Semiconductor Technology International Conference (CSTIC)*. 2020, pp. 1–4. DOI: 10.1109/CSTIC49141.2020.9282433.

[101] R. Ranica et al. "A capacitor-less DRAM cell on 75nm gate length, 16nm thin fully depleted SOI device for high density embedded memories". In: *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004.* 2004, pp. 277–280. DOI: 10.1109/IEDM.2004.1419131.

[102] E. Yoshida and T. Tanaka. "A capacitorless 1T-DRAM technology using gate-induced drain-leakage (GIDL) current for low-power and high-speed embedded memory". In: *IEEE Transactions on Electron Devices* 53.4 (2006), pp. 692–697. DOI: 10.1109/TED.2006.870283.

[103] Hsing-Jen Wann and Chenming Hu. "A capacitorless DRAM cell on SOI substrate". In: *Proceedings of IEEE International Electron Devices Meeting.* 1993, pp. 635–638. DOI: 10.1109/IEDM.1993.347280.

[104] Shan Deng et al. "A Comprehensive Model for Ferroelectric FET Capturing the Key Behaviors: Scalability, Variation, Stochasticity, and Accumulation". In: *2020 IEEE Symposium on VLSI Technology.* 2020, pp. 1–2. DOI: 10.1109/VLSITechnology18217.2020.9265014.

[105] Milan Pešić and Christopher Künneth. "A Computational Study of Hafnia based Ferroelectric Memories: From Ab-initio via Physical Modeling to Circuit Models of Ferroelectric Device". In: *Journal of Computational Electronics* 16.4 (2017), pp. 1236–1256.

[106] Jun-Han Hsia et al. "Dynamical Effects of Excess Carriers On SOI FeFET Memory Device Operations". Under review.

[107] R. R. Mehta, B. D. Silverman, and J. T. Jacobs. "Depolarization fields in thin ferroelectric films". In: *Journal of Applied Physics* 44.8 (1973), pp. 3379–3385. DOI: 10.1063/1.1662770. eprint: https://doi.org/10.1063/1.1662770. URL: https://doi.org/10.1063/1.1662770.

[108] Nanbo Gong and Tso-Ping Ma. "Why Is FE–HfO2 More Suitable Than PZT or SBT for Scaled Nonvolatile 1-T Memory Cell? A Retention Perspective". In: *IEEE Electron Device Letters* 37.9 (2016), pp. 1123–1126. DOI: 10.1109/LED.2016.2593627.

[109] AM Bratkovsky and AP Levanyuk. "Depolarizing field and "real" hysteresis loops in nanometer-scale ferroelectric films". In: *Applied physics letters* 89.25 (2006), p. 253108.

[110] Jae Young Kim, Min-Ju Choi, and Ho Won Jang. "Ferroelectric field effect transistors: Progress and perspective". In: *APL Materials* 9.2 (2021), p. 021102. DOI: 10.1063/5.0035515. eprint: https://doi.org/10.1063/5.0035515. URL: https://doi.org/10.1063/5.0035515.

[111] Jike Lyu et al. "High polarization, endurance and retention in sub-5 nm Hf0.5Zr0.5O2 films". In: *Nanoscale* 12 (20 2020), pp. 11280–11287. DOI: 10.1039/D0NR02204G. URL: http://dx.doi.org/10.1039/D0NR02204G.

[112] Sourabh Jindal et al. "Investigation of Gate-Length Scaling of Ferroelectric FET". In: *IEEE Transactions on Electron Devices* 68.3 (2021), pp. 1364–1368. DOI: 10.1109/TED.2021.3054720.

[113] Wenwu Xiao et al. "Performance Improvement of Hf0.5Zr0.5O2-Based Ferroelectric-Field-Effect Transistors With ZrO2 Seed Layers". In: *IEEE Electron Device Letters* 40.5 (2019), pp. 714–717. DOI: 10.1109/LED.2019.2903641.

[114] H. Mulaosmanovic et al. "Evidence of single domain switching in hafnium oxide based FeFETs: Enabler for multi-level FeFET memory cells". In: *2015 IEEE International Electron Devices Meeting (IEDM)*. 2015, pp. 26.8.1–26.8.3. DOI: 10.1109/IEDM.2015.7409777.

[115] Nujhat Tasneem et al. "The Impacts of Ferroelectric and Interfacial Layer Thicknesses on Ferroelectric FET Design". In: *IEEE Electron Device Letters* (2021), pp. 1–1. DOI: 10.1109/LED.2021.3088388.

[116] Philippe C. Adell et al. "Band-to-Band Tunneling (BBT) Induced Leakage Current Enhancement in Irradiated Fully Depleted SOI Devices". In: *IEEE Transactions on Nuclear Science* 54.6 (2007), pp. 2174–2180. DOI: 10.1109/TNS.2007.911419.

[117] Asif Islam Khan and Korok Chatterjee. "Negative capacitance in a ferroelectric capacitor". In: *Nature Materials* 14.2 (2015), pp. 182–186. DOI: 10.1038/nmat4148.

[118] Asif Islam Khan et al. "Understanding negative capacitance dynamics in ferroelectric capacitors". In: *2015 Fourth Berkeley Symposium on Energy Efficient Electronic Systems (E3S)*. 2015, pp. 1–3. DOI: 10.1109/E3S.2015.7336793.

[119] S. Sivasubramanian, A. Widom, and Y. Srivastava. "Equivalent circuit and simulations for the Landau-Khalatnikov model of ferroelectric hysteresis". In: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 50.8 (2003), pp. 950–957. DOI: 10.1109/TUFFC.2003.1226538.