

# **Cryogenic CMOS Compact Modeling for Cryo-Electronic Applications**

by

**Hazem Elgabra**

A thesis  
presented to the University of Waterloo  
in fulfillment of the  
thesis requirement for the degree of  
Doctor of Philosophy  
in  
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2023

© Hazem Elgabra 2023

### **Examining Committee Membership**

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:       Eva Dupont-Ferrier  
Associate Professor, Dept. of Physics and Institut quantique,  
Université de Sherbrooke

Supervisor:               Lan Wei  
Associate Professor, Dept. of Electrical and Computer Engineering,  
University of Waterloo

Internal Members:       Youngki Yoon  
Associate Professor, Dept. of Electrical and Computer Engineering,  
University of Waterloo

Guo-Xing Miao  
Associate Professor, Dept. of Electrical and Computer Engineering  
and Institute for Quantum Computing, University of Waterloo

Internal-External Member: Jonathan Baugh  
Professor, Dept. of Chemistry and Institute for Quantum Computing,  
University of Waterloo

### **Author's Declaration**

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Abstract

Quantum computing holds the promise of a monumental leap in computational power, enabling the resolution of previously insurmountable problems with astonishing speed compared to classical computers. Emerging computing paradigms, including Shor’s factoring algorithm, Grover’s searching algorithm, quantum simulations, protein folding, and more, stand on the brink of feasibility, thanks to quantum computers. However, despite impressive recent advancements in quantum computing, demonstrated systems featuring anywhere between a few to around a hundred physical qubits remain significantly distant from achieving quantum supremacy over classical computing, which demands the utilization of millions of physical qubits. This formidable challenge is known as the scaling problem.

Implementing large-scale quantum computing systems is faced with numerous hurdles, particularly where each qubit necessitates precise control under extremely low cryogenic temperatures ( $< 1$  K). Complementary metal-oxide-semiconductor (CMOS) technology, the cornerstone of classical computer scaling, emerges as a promising solution for scaling quantum computers. CMOS technology offers deep miniaturization and versatility, functioning seamlessly at both room temperature (RT) and cryogenic temperatures (cryoT). CMOS is compatible with the spin qubits in semiconductor quantum dots (one of the various methods of implementing qubits that exhibit long coherence time) offering integration compatibility especially from the fabrication perspective. It is this kind of tight integration that may ultimately hold the key to resolving the quantum scaling problem, bridging the gap between the current state of quantum computing and its promising potential.

Nevertheless, current circuit design environments lack support for operating temperatures near cryoT. This lack of support is centered in the often overlooked component known as the compact model. Compact models act as the blueprint that informs circuit simulators of how circuit elements behave under various operating conditions. This component is composed of simplified mathematical formulas that bridge the gap between the element’s physical model and simulation engines. In order to obtain accurate simulation results necessary for cryo-circuits design the compact model must be accurate. To obtain precise simulation results necessary for cryo-circuit design, it is imperative to understand and incorporate the effects of cryoT on metal-oxide-semiconductor field-effect-transistors (MOSFETs) into the compact model.

This thesis is one of the first attempts to develop cryogenic MOSFET compact model based on virtual source concept, through theoretical investigation and experimental validation. The cold temperature effects on MOSFETs are studied and integrated into the existing MIT virtual-source model (MVS), expanding its temperature range to include deep cryoT in the range of few Kelvin. To achieve this, sample devices from multiple commercial technology nodes are characterized

in RT down to deep cryoT. The thesis outlines the measurement setup and explores a range of predicted and unexpected cryogenic phenomena within the transistor. Good agreement between experimental data and modeled data is obtained between 300 K and 4 K for 20, 28, and 65 nm bulk CMOS technology nodes.

However, merely developing a cryogenic compact model is insufficient for its adoption and practical deployment. Extraction and fitting tools are therefore developed alongside the model. To support circuit design on industrial tools and validate the model, the compact model is implemented Verilog-A. Subsequently, a cryogenic circuit simulation is demonstrated using industry-standard electronic design automation (EDA) tools. This demonstration underscores the viability of the model to facilitate cryo-circuit design for quantum computing, representing a significant step towards realizing the potential of quantum computing in practical applications.

## Acknowledgements

This research is funded by Ontario Early Researcher Award, Quantum Quest Seed Fund (CFREF-TQT), the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, NSERC Canada Graduate Scholarship - Doctoral (CGS D), Postgraduate Scholarship - Doctoral (PGS D), and the University of Waterloo's Engineering Excellence Doctoral Fellowship.

Words are not enough to convey my appreciation and gratitude for my remarkable supervisor, Prof. Lan Wei. I am in ever-renewing awe of how Prof. Wei manages the research group, shuffles between many tasks, and yet maintain a just-right amount of macro and micro grasp and influence on the research directions and outcomes. Prof. Wei nurtured my writing, research, and critical thinking skills. More importantly, she gave me sufficient space to find my flow and research philosophy, and she trained me to be patient, agile, and resilient. I am wholeheartedly indebted to your guidance and support.

I wish to extend my deep gratitude and admiration to the members of the Coherent Spintronics Group. It was an enriching, collaborative experience for everything I learned from each member and for the pleasant experiences. Thank you, Francois Sfigakis and Brandon Buonacorsi for the countless hours spent in the cryogenic laboratory, even on weekends, and the good times we had together.

My sincere appreciation also goes to the internal thesis committee, Profs. Jonathan Baugh, Youngki Yoon, and Irene Goldthorpe, for devoting their valuable time and intellect to oversee my progress, and for the fruitful critiques and observations during the proposal. I am also very grateful to Prof. Guo-Xing Miao for graciously agreeing to join us as a replacement internal committee member and Prof. Eva Dupont-Ferrier for graciously agreeing to join us as an external examiner, and for setting aside her precious time and effort to examine the thesis and visit.

I wish to extend my deep gratitude and admiration to the members of the Waterloo Emerging Integrated Systems (WEIS) research group. It was an enlightening, cooperative experience for everything I learned from each member and for the wonderful memories. Thank you to Xuesong Chen, Rubaya Absar, Dylan Ma, Ryan Fang, Kaship Sheikh, Yiju Zhao, Jessica Chong, and Yijing Feng.

Going by Maslow's hierarchy of needs, this research would not have seen any light has it not been for feeling a sense of connection and love with my friends at the Waterloo region. Thank you Omar Alhussein, Kyle Booker, Chris Smith, Vivek T, Romina CT, Henry Ho, John Morgan, Hannah Zhang, Jordan Partridge, Helen Qin, Martin Zou, Manaf Ben Yeha, Jamal Busnaina, Ilyas Farhat, Omar Farhat, Jinane Harmouche, Hesham Moussa, Mohammad Shahab, Mohammad Shorman, Kirby Lau, and Leslie Zhang. Finally, I want to thank Jessica Gurd for her

ever-growing love, support, and patience. And, my deepest gratitude and affection goes to my parents, brothers, and sisters, for their unparalleled love! One call with them has metaphysical effects.

# Table of Contents

|                                                                                    |            |
|------------------------------------------------------------------------------------|------------|
| <b>Examining Committee Membership</b>                                              | <b>ii</b>  |
| <b>Author’s Declaration</b>                                                        | <b>iii</b> |
| <b>Abstract</b>                                                                    | <b>iv</b>  |
| <b>Acknowledgements</b>                                                            | <b>vi</b>  |
| <b>List of Abbreviations</b>                                                       | <b>xi</b>  |
| <b>List of Symbols</b>                                                             | <b>xiv</b> |
| <b>1 Introduction</b>                                                              | <b>1</b>   |
| 1.1 Motivation . . . . .                                                           | 1          |
| 1.2 Research Objectives . . . . .                                                  | 4          |
| 1.2.1 Cryogenic Device Characterization . . . . .                                  | 4          |
| 1.2.2 Semiconductor and Device Physics in CryoT . . . . .                          | 5          |
| 1.2.3 CryoT Compact Model . . . . .                                                | 6          |
| 1.2.4 Model Implementation, Parameter Extraction, and Circuit Simulation . . . . . | 6          |
| 1.3 Thesis Outline . . . . .                                                       | 7          |

|          |                                                      |           |
|----------|------------------------------------------------------|-----------|
| <b>2</b> | <b>Background</b>                                    | <b>9</b>  |
| 2.1      | MOSFET Devices . . . . .                             | 10        |
| 2.2      | MOSFET at CryoT . . . . .                            | 13        |
| 2.3      | Compact Models . . . . .                             | 17        |
| 2.4      | Progress in Cryogenic Compact Models . . . . .       | 20        |
| <b>3</b> | <b>Cryogenic Measurements</b>                        | <b>24</b> |
| 3.1      | Measurement Samples . . . . .                        | 25        |
| 3.1.1    | CMC65 Chip Design . . . . .                          | 26        |
| 3.2      | Cryogenic Measurement Setup . . . . .                | 31        |
| 3.2.1    | Dipper Probe . . . . .                               | 33        |
| 3.2.2    | Sample Preparation . . . . .                         | 34        |
| 3.2.3    | Experimental Setup . . . . .                         | 36        |
| 3.3      | Experimental Results . . . . .                       | 38        |
| 3.3.1    | Dipping Probe Characteristics . . . . .              | 38        |
| 3.3.2    | Device Characteristics . . . . .                     | 39        |
| <b>4</b> | <b>Cryogenic MOSFET</b>                              | <b>43</b> |
| 4.1      | Device Physics . . . . .                             | 43        |
| 4.1.1    | Field-Assisted Ionization . . . . .                  | 45        |
| 4.1.2    | Fermi Level Location . . . . .                       | 46        |
| 4.1.3    | Threshold Voltage . . . . .                          | 51        |
| 4.1.4    | Subthreshold Swing . . . . .                         | 55        |
| 4.1.5    | Carrier Transport . . . . .                          | 58        |
| 4.2      | The Compact Model . . . . .                          | 60        |
| 4.2.1    | Base Model: MIT Virtual Source Model (MVS) . . . . . | 60        |
| 4.2.2    | The Cryogenic MVS model . . . . .                    | 62        |

|          |                                                        |            |
|----------|--------------------------------------------------------|------------|
| <b>5</b> | <b>Readiness for Practical Simulations</b>             | <b>73</b>  |
| 5.1      | Implementation . . . . .                               | 73         |
| 5.1.1    | Verilog-A for Compact Models . . . . .                 | 74         |
| 5.1.2    | MVS model Verilog-A implementation . . . . .           | 76         |
| 5.2      | Supporting tools . . . . .                             | 80         |
| 5.2.1    | Parameters Extraction . . . . .                        | 82         |
| 5.2.2    | Single Temperature Extraction . . . . .                | 83         |
| 5.2.3    | Temperature Dependence Extraction . . . . .            | 91         |
| 5.3      | Circuit Simulation . . . . .                           | 96         |
| <b>6</b> | <b>Conclusions and Future Research</b>                 | <b>99</b>  |
| 6.1      | Conclusions . . . . .                                  | 99         |
| 6.2      | Contributions . . . . .                                | 100        |
| 6.3      | Future Research . . . . .                              | 101        |
|          | <b>References</b>                                      | <b>102</b> |
|          | <b>APPENDICES</b>                                      | <b>114</b> |
| <b>A</b> | <b>Analytical derivation for Fermi energy location</b> | <b>115</b> |
| <b>B</b> | <b>Analytical Derivation for Fermi Potential</b>       | <b>117</b> |
| B.1      | Incomplete Ionization Case: . . . . .                  | 117        |
| B.2      | Complete Ionization Case: . . . . .                    | 118        |
| <b>C</b> | <b>Verilog-A Code for CryoMVS Model</b>                | <b>119</b> |
|          | <b>Glossary</b>                                        | <b>134</b> |

# List of Abbreviations

**ADC** analog-to-digital converter 21

**BJT** Bipolar Junction Transistor 28, 30, 31, 37

**CM** compact model 20–23, 43, 73–78, 80, 81, 96, 100–102

**CMC** Compact Model Council 75–77, 79, 80

**CMOS** Complimentary Metal-Oxide-Semiconductor 3, 4, 7, 9–11, 15, 20, 99–101

**cryoMVS** cryogenic MVS 62, 63, 65–71, 74, 76, 77, 79, 81, 93, 96, 101, 102

**Demux** Demultiplexer circuit 23, 96, 100

**DIBL** Drain-Induced-Barrier-Lowering xv, 61, 62, 85, 88, 89

**DoS** density of states 56, 57, 63, 64

**DUT** device under test 5, 24, 31, 32, 35–38, 81, 82, 93, 101

**EDA** electronic design automation 6, 7, 75, 96, 99, 100

**ESD** electrostatic discharge 26–29, 36, 40, 99

**FD** Fermi-Dirac 44, 47, 48

**FDSOI** Fully-Depleted SOI 15, 59, 101

**FET** Field-Effect Transistor xii, 9, 10, 14, 45, 58, 59, 61, 65, 79

**GAA** Gate All-Around 9

**GF** GlobalFoundries 18

**GUI** graphical user interface 7, 82, 85–87, 89–95, 102

**I/O** input/output 26–29

**IC** integrated circuit 17, 18

**IQC** Institute for Quantum Computing 5, 25

**LNA** low-noise amplifier 21, 22

**MB** Maxwell-Boltzmann 47, 48, 50, 100

**MOS** Metal-Oxide-Semiconductor xi, xii, xv, xvii, xix, 10, 51, 52

**MOSFET** Metal-Oxide-Semiconductor Field-Effect Transistor xii, xix, 7, 9–15, 17, 18, 21, 22, 43, 51, 61, 99, 101, 134

**MVS** MIT Virtual-Source xi, 20–23, 60–66, 76, 79, 84, 86, 88–90, 93, 94, 100, 101

**NMOS** n-type MOSFET 11, 14, 19, 28, 41, 52, 62, 66–69, 71, 72, 90

**OSC** oscillator 21

**PDK** process design kit 18, 19, 27, 99

**PMOS** p-type MOSFET 11, 28, 42, 62, 66–69, 71, 72, 90

**QD** quantum dot 10

**QEC** quantum error correction 2, 96

**QNFCF** Quantum-Nano Fabrication and Characterization Facility 5

**QPU** quantum processing unit 9

**qubit** quantum bit 1–4, 14

**RAC** Research Advancement Center 5, 25, 31, 37

**RO** ring-oscillator 21, 22, 89

**SA** Spanish Acquisition 37

**SI** the international system of units 76

**Si** silicon xv, xviii, 14, 15, 47–49, 53, 54, 56

**SMU** source-measure unit 37, 38, 99

**SOI** Silicon-on-Oxide xi, 61, 101

**SPICE** Simulation Program with Integrated Circuit Emphasis 18

**TSMC** Taiwan Semiconductor Manufacturing Company 18, 19

**VA** Verilog-A 7, 74–77, 79, 80, 100

**VS** Virtual-Source xii, 60, 61, 65

# List of Symbols

- $\Phi_{MS}$  The metal-semiconductor work-function defined as:  $\Phi_M - \Phi_S$  in [eV] [52](#)
- $\Phi_M$  The metal work-function defined as the energy difference between  $E_0$  and  $E_F$  in the metal [eV] [xiv](#), [52](#), [53](#)
- $\Phi_S$  The semiconductor work-function defined as the energy difference between  $E_0$  and  $E_F$  in the semiconductor in [eV] [xiv](#), [52–54](#)
- $Q_D$  Depletion charge [53](#), [54](#)
- $Q_{it}$  Interface trap charge [64](#)
- $Q_{ix_o}$  Charge density at the virtual source, or top of the barrier, in  $[m^{-3}]$  [60](#), [61](#)
- $Q_o$  Localized charge density near the band gap edges [57](#)
- $Q_s$  Semiconductor charge including mobile charge that contribute to current, and localized charge that does not contribute to current [52](#), [53](#)
- $\chi$  The electron affinity defined as the energy difference between  $E_0$  and  $E_C$  in [eV] [52](#), [53](#)
- $a_1$  A fitting parameter that describes the magnitude of  $\mu_{app}$  below  $T_{mu}$  [63](#), [66](#), [72](#), [94–96](#)
- $a_2$  A fitting parameter that describes the magnitude of  $\mu_{app}$  above  $T_{mu}$  [63](#), [66](#), [72](#), [94–96](#)
- $\alpha$  Empirical fitting parameter used for transitioning between weak and strong inversion. [Unitless] [xiv](#), [61](#), [65](#), [66](#)
- $b_1$  A fitting parameter that describes the magnitude of  $v_{inj}$  below  $T_{vxo}$  [63](#), [66](#), [72](#), [95](#)
- $b_2$  A fitting parameter that describes the magnitude of  $v_{inj}$  above  $T_{vxo}$  [63](#), [66](#), [72](#), [95](#)

- $\beta$  Empirical fitting parameter used for transitioning between linear and saturation regions. [Unitless] 61, 62, 89
- $C_{inv}$  Inversion capacitance per area in  $[F/m^2]$  61
- $C_{ox}$  Oxide capacitance per area in  $[F/m^2]$  xviii, 52, 54, 57, 65
- $c_{V_{th}}$  A fitting parameter that holds the physical meaning of dopant concentration ( $N_A$  or  $N_D$ ) and the  $T$ -independent portion of  $N_V$  and  $N_C$  62, 63, 72, 94
- $\delta$  DIBL factor. [Unitless] 61, 86–89
- $D_{it}$  Density of interface traps xv–xviii, 22, 55, 56, 64, 65
- $D_u$  Uniformly distributed  $D_{it}$  xvi, 57, 65
- $E_0$  The vacuum level. The minimum level an electron must gain to completely free itself from the material in [eV] xiv, 52, 53
- $E_A$  Acceptor energy level in [eV] xvii, 45–48, 50, 52, 53, 62, 115, 116
- $E_C$  Energy level at the bottom of the conduction band (or the top of the band gap) in [eV] xiv, 45, 47, 48, 50, 56, 57
- $E_D$  Donor energy level in [eV] 45, 47, 48, 50
- $E_F$  The Fermi level or Fermi energy in [eV]. The energy level where the Fermi function equals 1/2. xiv, xv, xvii, 44–53, 55, 56, 64, 100, 115, 116
- $E_G$  Band gap energy, or the gap in energy between the bound states (valance band) and the free states (conduction band). In [eV] 45, 48, 49, 52, 62
- $E_i(y)$  The intrinsic energy along the vertical dimension of a MOS device in [eV] xvii, 51
- $E_i$  Intrinsic Fermi level, or  $E_F$  when  $n = p$ , in [eV] xvii, 45, 51–53
- $\epsilon_{Si}$  Silicon permittivity valued at [1]:  $11.7 \times \epsilon_0 = 1.036 \times 10^{-12}$  in [F/cm] xviii, 53, 54
- $E_V$  Energy level at the top of the valance band (or the bottom of the band gap) in [eV] 45, 47, 48, 50, 52, 57, 62, 115, 116
- $E$  Energy in [eV] xvi, 44

- $f(E)$  Fermi-Dirac distribution function with respect to  $E$ . Typically used to determine the probability of occupation, and occasionally, as a window or smoothing function. [Unitless] 44, 57
- $F_f$  Fermi transition function for transitioning between weak and strong inversion from the output transfer perspective. [Unitless] xix, 61, 65, 66
- $F_S$  Saturation function for transitioning between linear and saturation regions from the output characteristics perspective. [Unitless] 61
- $g_A$  Acceptor degeneracy factor, not related to semiconductor degeneracy, or the number of carriers an acceptor cite can accommodate. The standard value for  $g_A$  is 4. [unitless] xvi, 47–50, 55, 62, 115–118
- $g_D$  Donor degeneracy factor, not related to semiconductor degeneracy, or the number of carriers a donor cite can accommodate. The standard value for  $g_D$  is 2. [unitless] xvi, 47–50, 115
- $I_D$  Drain current in [A] 11, 12, 14–16, 40–42, 61, 79, 82, 84, 85, 88–91, 97
- $k_B$  Boltzman constant equal to 1.380649e-23 [J/K] 14, 15, 44, 46–50, 56, 62, 65, 77, 115, 116
- $L_{eff}$  Effective channel length with overlap on the source and drain sides removed in [m] 61, 62, 89
- $L$  Transistor length in, gate length, or channel length in [m] 13
- $m_0$  Subthreshold factor at RT. [Unitless] 56, 61, 62, 65, 84–86, 90, 93
- $m_d$   $V'_{DS}$ -dependence portion of the subthreshold factor. [Unitless] 61, 62, 65, 84–86, 88, 93
- $m_{exp}$  Subthreshold factor due to exponential  $D_{it}$ . [Unitless] 64, 65
- $m_{it}$  Subthreshold factor due to  $D_{it}$ . [Unitless] 65
- $m_{S0}$  Subthreshold factor due to  $D_u$ . [Unitless] 64, 65
- $\mu_{app}$  Apparent mobility in  $[m^2/V.s]$  xiv, xvii, xviii, 59, 65, 89, 90, 94, 95, 100
- $\mu_{bal}$  Ballistic mobility in  $[m^2/V.s]$  59
- $\mu_{eff}$  Effective mobility in  $[m^2/V.s]$  21, 58
- $\mu$  Mobility in  $[m^2/V.S]$  6, 13, 14, 16, 21, 22, 58, 59, 61, 63, 66, 69, 72, 91, 94, 96, 100

- $m$  Total subthreshold factor. [Unitless] 14, 15, 55–58, 61–65, 84, 85, 88, 93
- $N_A^-$  Ionized acceptor atom concentration in  $[cm^{-3}]$  46, 47, 115, 117
- $N_A$  Acceptor atom concentration in  $[cm^{-3}]$  xv, xviii, 47, 48, 50, 53, 55, 63, 115–118
- $N_C$  Effective density of conduction band states in  $[cm^{-3}]$  xv, 47, 48, 50, 62, 63, 72
- $N_D^+$  Ionized donor atom concentration in  $[cm^{-3}]$  46–48, 50
- $N_D$  Donor atom concentration in  $[cm^{-3}]$  xv, xviii, 47–50
- $n_i$  The intrinsic carrier concentration in  $[cm^{-3}]$  49, 55, 62, 117, 118
- $N_o$  A fitting parameters related to the magnitude of the exponential portion of  $D_{it}$  63, 65, 93
- $N_V$  Effective density of valance band states in  $[cm^{-3}]$  xv, 47, 48, 50, 62, 63, 72, 115, 116
- $n$  Electron concentration in  $[cm^{-3}]$  xv, 20, 46–50, 72
- $p_1$  A fitting parameter that describes the power-law effect of  $T$  on  $\mu_{app}$  below  $T_{mu}$  63, 66, 72, 94–96
- $p_2$  A fitting parameter that describes the power-law effect of  $T$  on  $\mu_{app}$  above  $T_{mu}$  63, 66, 72, 94–96
- $\phi_A$  The acceptor ionization potential defined as a constant anywhere in the semiconductor:  $\frac{E_i - E_A}{q}$  in [V] 52, 55, 117, 118
- $\phi_F$  The Fermi potential defined at the bulk as:  $\frac{E_i(bulk) - E_F}{q}$  in [V] 51, 52, 54, 55, 62, 86, 117, 118
- $\phi_s$  The surface potential defined at the semiconductor-oxide interface:  $\frac{E_i(bulk) - E_i(0)}{q}$  in [V] 14, 51–54, 57, 65
- $\phi_T$  Thermal potential or voltage in [V] xix, 14, 55, 56, 58, 61–63, 65, 66, 84, 117, 118
- $\phi(y)$  The internal potential along the vertical dimension of a MOS device defined as:  $\frac{E_i(bulk) - E_i(y)}{q}$  in [V] 51
- $\phi$  Potential. Similar to voltage, but often used for semiconductor related potentials in [V] 51
- $p$  Hole concentration in  $[cm^{-3}]$  xv, 46–48, 72, 115–117

$q_1$  A fitting parameter that describes the power-law effect of  $T$  on  $v_{inj}$  below  $T_{vxo}$  63, 66, 72, 95  
 $q_2$  A fitting parameter that describes the power-law effect of  $T$  on  $v_{inj}$  above  $T_{vxo}$  63, 66, 72, 95  
 $q$  Elementary charge equal to 1.602176634e-19 J or C, or 1 eV xvii, xviii, 14, 15, 51–53, 57, 77  
 $R_0$  Internal resistance in  $[\Omega.m]$  88, 89, 91  
 $R_D$  Internal drain resistance in  $[\Omega]$  60, 88  
 $\rho$  Total charge density in  $[cm^{-3}]$  46  
 $R_S$  Internal source resistance in  $[\Omega]$  60, 88  
 $S_0$   $S$  offset parameter in [mV/dec] 63, 65, 72, 92  
 $sm_{mu}$  Smoothing factor in the window function of  $\mu_{app}$  model. [Unitless] 63, 95  
 $sm^*$   $V_{gs}^*$  smoothing factor. [Unitless] 65  
 $sm_{vxo}$  Smoothing factor in the window function of  $v_{inj}$  model. [Unitless] 63, 95  
 $sm$   $S$  saturation smoothing factor. [Unitless] 63, 64, 72, 92  
 $SS$  subthreshold slope in [dec/mV] 12  
 $S$  Subthreshold swing in [mV/dec] xviii, 6, 12–15, 20–22, 40–42, 55–58, 63–65, 68, 72, 84, 92, 93, 100  
 $T_e$  Effective temperature, related to the  $S$  saturation effect, in [K] 64, 65  
 $\theta$  A fitting parameter that holds the physical meaning of  $q$ , dopant concentration ( $N_A$  or  $N_D$ ), the  $\epsilon_{Si}$ , and  $C_{ox}$  54, 56, 63, 94  
 $T_{mu}$  The critical temperature in the  $\mu_{app}$  model in [K] xiv, xvii, 63, 66, 72, 94, 95  
 $T_o$  A fitting parameters related to the decay energy of the exponential portion of  $D_{it}$  63, 65, 93  
 $T_S$   $S$  saturation temperature in [K] 56, 63, 64, 66, 72, 92, 93  
 $T_{vxo}$  The critical temperature in the  $v_{inj}$  model in [K] xiv, xviii, 63, 66, 72, 94, 95  
 $T$  Temperature in [K] xv, xvii, xviii, 14, 15, 21, 22, 44–50, 53–59, 62–69, 82, 91–96, 98, 115, 116

$v_{inj}$  Injection velocity in [m/s] [xiv](#), [xviii](#), [59](#), [65](#), [94–96](#), [100](#)  
 $V_{DSAT}$  Saturation voltage in [V] [12](#), [61](#), [62](#), [85](#), [89](#), [90](#)  
 $V_{DS,lin}$  Drain-source voltage in the linear region in [V] [84–86](#), [88](#)  
 $V'_{DS}$  Drain-source voltage in [V] [xvi](#), [xix](#), [60–62](#), [65](#), [85](#), [88](#), [93](#)  
 $V_{DS}$  Drain-source voltage in [V] [10–13](#), [15](#), [16](#), [35](#), [36](#), [39–42](#), [61](#), [70](#), [79](#), [88–93](#)  
 $V_D$  Drain voltage in [V] [82](#), [84](#), [85](#), [88](#), [89](#), [97](#)  
 $V'_{GS}$  Internal gate-source voltage in [V] [xix](#), [60](#), [61](#), [65](#)  
 $V_{gs}^*$  Saturated  $V'_{GS}$  in [V] [xviii](#), [65](#)  
 $V_{GS}$  Gate-source voltage in [V] [10–15](#), [20–22](#), [35](#), [39–42](#), [55](#), [57](#), [64](#), [70](#), [77](#), [80](#), [82](#), [88–91](#), [93](#), [100](#)  
 $V_G$  Gate voltage in [V] [14](#), [51–53](#), [56](#), [63](#), [64](#), [79](#), [80](#), [82](#), [84](#), [97](#)  
 $V_r$  Reference voltage equal to  $V_{th} - \alpha\phi_T F_f/2$ . In [V] [65](#)  
 $V_{T0}$   $V_{th}$  at  $V'_{DS} = V'_{GS} = 0$ . In [V] [61–63](#), [67](#), [86–88](#), [90](#), [91](#), [93](#)  
 $V_{th}$  Threshold voltage in [V] [xix](#), [6](#), [12](#), [14–16](#), [20–22](#), [52–54](#), [58](#), [61–63](#), [65](#), [77](#), [85–88](#), [93](#), [94](#), [98](#), [100](#)  
 $v_T$  Thermal velocity in [m/s] [16](#), [59](#)  
 $v_{xo}$  Local carrier velocity at the virtual source, or top of the barrier, in [ $m/s$ ] [60](#), [61](#), [63](#), [66](#), [69](#), [89–91](#), [94](#)  
 $v$  Velocity in [m/s] [6](#), [13](#), [14](#), [58](#), [59](#), [100](#)  
 $W$  Transistor width in [m] [61](#), [88](#)  
 $y$  Vertical spacial dimension in a MOSFET or MOS device in [m] [52](#), [53](#)

# Chapter 1

## Introduction

### 1.1 Motivation

The emerging paradigm of quantum computing is ushering in a new era of computing, with potential for massive computational power that classical computers cannot achieve. By harnessing quantum mechanics phenomena such as quantum entanglement and superposition, quantum computers have the ability to solve intractable problems exponentially faster than classical supercomputers [2, 3]. This capability arises from the unique properties of quantum bits (qubits), which can exist in a superposition of both 0 and 1 and can simultaneously hold a portion of each state. This allows a quantum processor with  $n$  qubits to perform computations on all possible  $2^n$  combinations simultaneously, as opposed to the one-at-a-time approach of classical computers. For example, A quantum processor with 2 qubits ( $n = 2$ ) performs only one computation instead of having to repeat a computation for each permutation (in this case: 00, 01, 10, 11) as done in classical computing [4]. This exponential increase in computational power has enormous potential in various fields including: cryptography, mathematics, chemistry, medicine, and material science. The computational opportunities offered by quantum computing include Shor's factoring algorithm, Grover's searching algorithm, quantum machine learning, and complex optimization, as well as quantum simulations for the synthesis of drugs and materials [3, 4].

Despite the unprecedented potential of quantum computing, this emerging technology is still in its infancy, with development currently reminiscent of that of the classical computer in the 1960s, in that it is incredibly resource and space intensive. While quantum computers can theoretically solve problems exponentially faster than classical supercomputers, achieving quantum supremacy requires  $10^2$  qubits [6]. However, to run practical quantum algorithms, the number of required qubits is projected to be in the order of millions. In 2019, Google claimed to have



Figure 1.1: Google’s Sycamore processor. Blue coaxial cables are connected from custom analog control electronics (on the right) to the quantum processor. Source: [5]

achieved quantum supremacy with its 53-qubit processor [6], but this claim has been challenged by other major players in the field like IBM [7]. Nevertheless, the sheer amount of wiring required for just 53 qubits, as displayed in Fig. 1.1, reveals the scalability challenge that quantum computing faces [5]. Moreover, achieving and sustaining high-fidelity qubits is very challenging. To address this issue, researchers have developed surface code techniques, including quantum error correction (QEC), that incorporate extra qubits. While this approach helps maintain accuracy, it also introduces significant redundancy, with approximately  $\sim 10^4$  physical qubits needed for each logical qubit. This approach exacerbates the scalability problem and amplifies the need for a viable scaling solution.

A quantum computer consists not only of a processor containing qubits, but also of interface components that enable the transmission of control signals and read-out of the qubits’ output. These interface components are based on classical electronics. Due to their extreme fragility, qubits are susceptible to noise that can compromise their delicate quantum state. To suppress thermal noise and maintain their quantum state, it is necessary to operate at extremely cold temperatures as close to absolute zero as possible (typically in the range of few mK). This is achieved through the use of sophisticated dilution fridges, such as the one illustrated in Fig. 1.2, which comprises multiple cooling stages. These fridges utilize complex cooling methods with different helium isotopes and are also shielded from all types of radiation in a copper room to reach cryogenic temperatures approaching absolute zero [8]. Interfacing with a quantum processor is

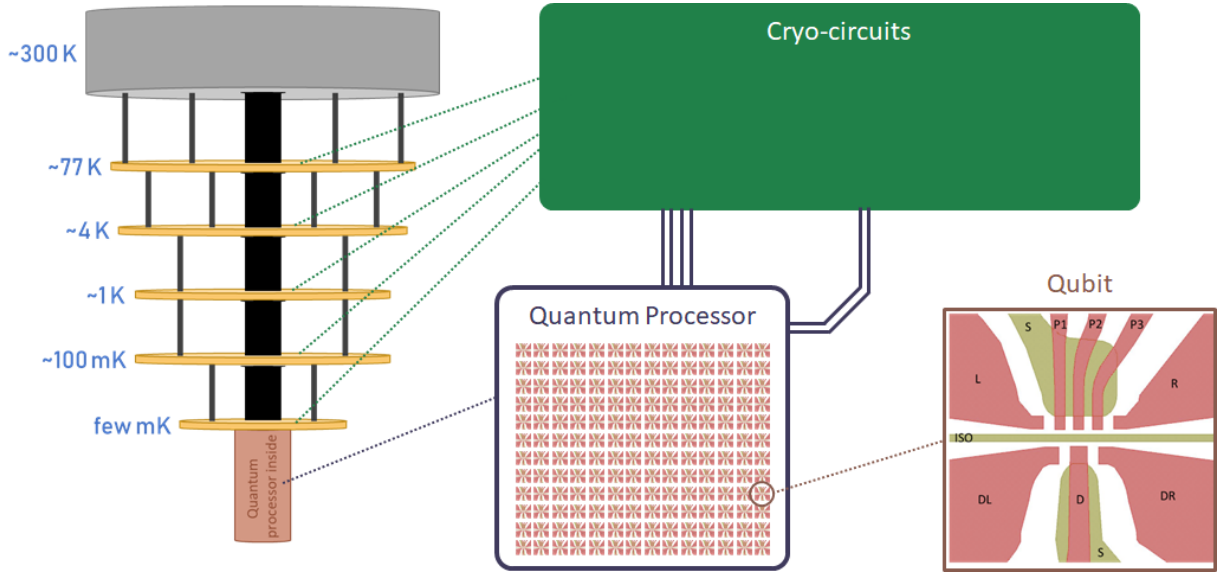


Figure 1.2: Quantum computer diagram showing a typical dilution fridge in which the computer resides. The quantum processor comprised of an array of qubits resides at the coolest stage while the classical control circuitry can be distributed in all stages.

a complex task that requires careful consideration of multiple factors.

As displayed in Fig. 1.1, researchers typically use shielded wires to connect laboratory instruments operating at room temperature (RT) to the quantum processor at the bottom layer of the dilution fridge operating at cryoT [9, 10]. However, as the number of qubits and shielded wires grows into the thousands and millions, this approach becomes increasingly impractical in terms of cost, reliability, and spacial requirements [3]. The associated direct coupling to RT and added noise can compromise both the cooling capability of the fridge and the integrity of the signals and qubit states. Therefore, utilizing cryo-circuits is the key to solving the quantum computer scaling problem by bringing the classical controller closer to the qubits, eliminating RT coupling and allowing researchers to focus on advancing quantum processors further. To achieve this, cryo-electronics can be housed at different cooling stages (as depicted in Fig. 1.2) depending on their functionality, power dissipation levels, and available cooling power in each stage.

While electronic technologies can operate at cryogenic temperatures (cryoT), Complimentary Metal-Oxide-Semiconductor (CMOS) technology offers the most promising scalable solutions due to their large-scale integration potential using established semiconductor processes. Standard CMOS is deeply scaled down in size and can operate at both RT and cryoT, making it an ideal candidate for complex system-on-chip quantum computers with thousands and millions of qubits

in the long run [9–12]. This potential has inspired a substantial research effort to achieve qubit-CMOS co-integration [13–19], as well as extensive studies on cryogenic CMOS devices and circuits [9–12, 20–32].

Qubits are very delicate devices that require precision in the input control signals and the ability to amplify weak output signals, while still maintaining the qubits’ quantum state. Circuit designers tasked with developing these highly specialized circuits are faced with a multitude of trade-offs while being constrained by stringent power and noise margins. Circuit designers typically use circuit simulators to optimize their designs which rely on compact device models to predict the behavior of their circuit. Compact models are a set of mathematical formulas that accurately describe the transistor’s behavior and predict their performance under certain operating conditions, include the operating temperature. Although CMOS devices and circuits are demonstrated to operate at cryogenic temperatures, their performance varies significantly from room temperature operation. Device models in the industry are typically developed for operating temperatures between  $-50^{\circ}\text{C}$  to  $200^{\circ}\text{C}$  (or  $223.15\text{ K}$  to  $473.15\text{ K}$ ). The objective of this project is to bridge the gap and enable circuit designers to design and optimize cryo-circuits suitable for scalable solid-state quantum processors with integrated cryo-electronics.

## **1.2 Research Objectives**

To bridge the temperature gap and facilitate the utilization of commercial technologies in cryogenic applications, this research encompasses four key objectives: cryogenic device characterization, exploration of semiconductor and device physics in cryoT, compact modeling, and implementation. Through the pursuit of these avenues, the groundwork will be laid for the creation of specialized cryo-electronics design systems tailored specifically for cryogenic applications, particularly in the field of quantum computing. The subsequent subsections will expand on each objective and provide detailed insights.

### **1.2.1 Cryogenic Device Characterization**

Physical models used in commercial technologies contain temperature parameters that may be used to predict device performance at cryoT. However, these models are usually only validated down to  $-50^{\circ}\text{C}$  (or  $223\text{ K}$ ), more than 100 degrees above the cryoT regime. Thus, the accuracy of these predictions must be evaluated, making it necessary to collect experimental data in the real world beyond the scope of simulation and modeling.

However, cryogenic measurements are non-trivial and several questions must be answered regarding the measurement setup, and the equipment needed. With the device under test (DUT) residing at temperature below 77 K (or  $-196^{\circ}\text{C}$ ) far from safe human reach, are there any considerations added by the equipment enabling such measurements? What are the challenges that such extreme temperature entail? Are there any additional instruments that may affect the measurement integrity and accuracy? Are the measurements required feasible within a reasonable time frame and cost?

With all these questions answered, it is also important to make an adequate measurement plan and clearly define the desired tests to be performed on which sample. Furthermore, there will be specialized cryo-components that are not available in the off-the-shelf market, and need to be designed and implemented specifically for this experiment. Chapter 3 discusses the design and implementation of an on-chip temperature sensor, and special custom chip holder. The chapter also details the challenges faced, including wire-bonding and the extensive time required to conduct a single measurement.

Achieving this objective was only possible through the collaboration with the quantum research group lead by Prof. Johnathan Baugh in the Institute for Quantum Computing (IQC) at the University of Waterloo. The cryogenic measurements were conducted using the cryogenic facilities in the Research Advancement Center (RAC) while samples were prepared using the Quantum-Nano Fabrication and Characterization Facility (QNFCF).

## 1.2.2 Semiconductor and Device Physics in CryoT

This objective primarily focuses on addressing the fundamental question: *Does our current understanding of semiconductor physics provide sufficient insights to accurately predict and model cryoT effects in MOSFET devices?*

With experimental data collected, the validity and accuracy of classical theories can be assessed in these extreme temperatures. This validation process becomes crucial to ascertain the applicability of established theories and models in cryogenic environments. It is especially vital to re-evaluate and confirm several common assumptions and practices regularly used in the field of semiconductor physics in cryogenic temperatures. For example, the assumption of complete ionization and the utilization of Maxwell-Boltzman statistics should be scrutinized and verified under cryoT conditions. Moreover, comparing experimental data with established theories will facilitate the identification of potential deviations and anomalies that may emerge exclusively at extremely low temperatures. Consequently, the development and modeling of a new theoretical framework may be necessary to comprehensively understand and explain these novel phenomena.

### 1.2.3 CryoT Compact Model

The core objective of this thesis is to develop a compact model that can accurately describe transistor performance down to cryoT, without compromising its RT capabilities. This model will also focus on key device performance parameters instead of overall device operation physics. These parameters include: threshold voltage ( $V_{th}$ ), subthreshold swing ( $S$ ), carrier mobility ( $\mu$ ), and velocity ( $v$ ). Compact models are of utmost importance to circuit designers, enabling them to predict the overall circuit behaviour under various operating conditions and circuit topologies comprised of countless devices. Given the importance of computational simplicity for large-scale circuit simulations, a physics-first semi-empirical approach is adopted with an emphasis on computational efficiency, low numbers of fitting parameters, and the flexibility to fit many processes.

While based on physics at their core, compact models have further sets of challenges than cryo-physics. They are also heavily affected by the principles of numerical and applied mathematics, and thus, must strike a balance between math and physics to produce a useful tool for circuit designers. Models are used as part of circuit simulators to describe the behaviour of thousands (if not millions) of transistors all at once. For this reason, these models need to be adequately optimized to use minimal computing power and must resolve in short time. The complex mathematical formulas used in physical models cannot be used as-is in compact models. Moreover, not all physical phenomena are included, and they are often simplified to provide the most relevant information for the expected operating conditions. Essentially, compact models require careful and accurate unification of semiconductor physics and applied mathematics to be useful.

### 1.2.4 Model Implementation, Parameter Extraction, and Circuit Simulation

A compact model alone is not user friendly without parameter extraction, model fitting tools, and of course, proper documentation. This requirement is addressed by developing specific software that can perform these tasks at RT down to deep cryoT. Additionally, the model must be made compatible with circuit design tools and simulators, and conform to modern compact modeling standards. As such, special considerations must be given to implementing complex mathematical expressions in Verilog-A to ensure convergence and quick simulation. Finally, successful cryogenic-circuit simulation must be achieved using industry-standard electronic design automation (EDA) tools, such as Cadence, Synopsys, Siemens, Silvaco, etc, to demonstrate the viability of cryo-circuits for quantum computing.

The research efforts under this objective will produce the following deliverables, allowing the model to be easily adopted by the research community:

- Matlab interface with graphical user interface (GUI) for:
  - Experimental data visualization and preparation.
  - Device parameters extraction.
  - Model fitting (or parameters fine-tuning).
- A Verilog-A (.va) version of the model, compatible with EDA tools.
- Circuit simulation of sample cryo-circuits for quantum computing.

## 1.3 Thesis Outline

The rest of this thesis is organized as follows. Chapter 2 provides the necessary background information on transistors and compact models. It starts with a brief overview of the advancements and current state of the semiconductor industry, specifically CMOS. Then, section 2.1 introduces device operation and defines relevant performance parameters, which is followed by section 2.2 introducing cryoT operation and its effect on device operation. After that, compact modeling is briefly introduced along with its importance and requirements (section 2.3), before shedding light on the state-of-the-art on cryogenic compact models (section 2.4). Next, chapters 3, 4, and 5 discuss how the aforementioned objectives are achieved in detail.

Chapter 3 focuses on cryogenic measurements, samples, setup, and results. Sample chips, including the one specifically designed for these measurements, are introduced in section 3.1. The purpose of each chip and included structures are presented. Then, section 3.2 describes the two measurement setups used throughout the study along with their challenges, advantages, and disadvantages. The chapter concludes with experimental results illustrated in section 3.3.

Cryogenic physics and device modeling are discussed in Chapter 4. Section 4.1 focuses the physical model in light of cryogenic operation, while section 4.2 converts it into a practical compact model. The latter section introduces the chosen base Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) compact model in subsection 4.2.1 and then details the cryogenic effects integration in subsection 4.2.2.

In Chapter 5, the cryoCMOS model is prepared for practical use in industry circuit design environments. First, section 5.1 talks about the language used for implementation, Verilog-A

(VA), its importance and popularity in today's chip design industry. The section also goes over the considerations and guidelines for writing Verilog-A models. Second, section 5.2 presents the developed software application for data preparation, parameter extraction, and model fitting along with automation capabilities. Lastly, a cryo-circuit designed for quantum computing application is presented and simulated both in RT and cryoT, showcasing the use of cryoMVS in practice.

Finally, concluding remarks are drawn in Chapter 6 along with suggestions for future research directions.

# Chapter 2

## Background

The silicon industry has made remarkable strides, propelling the world into an era of highly advanced electronic devices. Central to this progress is the development of the Complementary Metal-Oxide-Semiconductor (CMOS) technology, which has served as the backbone of modern electronics. Over the years, the semiconductor industry has relentlessly pursued Moore's Law, a driving force that has led to exponential improvements in the development of Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET), the building block of CMOS. This drive to scale down transistor dimensions and push the limits of what is technologically feasible has brought forth a multitude of advancements, from increased transistor density to enhanced performance and energy efficiency. The cutting-edge CMOS technologies (such as FinFET and Gate All-Around (GAA)-Field-Effect Transistors (FETs)) have achieved unprecedented levels of integration and computational power, catering to the demanding requirements of various applications, including high-performance computing, mobile devices, and artificial intelligence. Fortunately, MOSFETs maintain their operation at cryoT (a temperature range they are not designed for) making it possible to bring this thriving infrastructure along with its design principles into cryo-electronics.

However, the technological advancements achieved have exceeded the specific needs of certain applications. Thus, depending on the application, older and more mature CMOS process nodes may present a more attractive solution than modern nodes at the forefront of the industry. In the realm of quantum computers, where the computing power primarily stems from the quantum processing unit (QPU), cryo-circuits are left with less performance-demanding tasks like interface and control. Since the focus is to achieve monolithic integration of the QPU and cryo-electronics on a single chip, it necessitates flexible layout rules and simpler technologies that are easier to comprehend, modify, and employ. Cost and yield also emerge as critical considerations, particularly during the research and development phase, where affordable and reliable chips can

help maintaining the focus on quantum dot (QD) development. Mature technologies such as Bulk MOSFET (although not at the cutting edge) have been meticulously honed and refined over the decades after their inception. Therefore, they offer cost-effective and scaled-down solutions for cryogenic CMOS applications with adequate performance and reliable yields. Furthermore, the maturity of these technologies ensures a broader knowledge base, as they have been widely accessible to engineers, allowing a larger talent pool of circuit designers compared to the limited accessibility of cutting-edge technologies within the silicon industry.

The following sections will focus on the typical MOSFET operation and highlight its performance parameters. Subsequently, the effects of cryogenic temperatures on MOSFET operation, per experimental observations, is explored. Then, transistor compact models will be introduced along with their crucial role in cryo-electronics design process and distinction from physical models. Finally, the state-of-the-art of cryogenic MOSFET models in literature will be explored.

*This thesis will use the words "transistor" and "device" interchangeably to refer to FET and its variations, such as MOSFET. This is not to be confused with complete electronic systems (such as smartphone and other gadgets) or circuits made of multiple FETs, which will be referred to as "electronics" instead.*

## 2.1 MOSFET Devices

MOSFET devices are fundamental building blocks of modern electronic systems, offering excellent control over the flow of current. At its core, a MOSFET consists of a silicon substrate with a thin insulating layer, known as the gate oxide, and a gate electrode on top (Fig. 2.1a). This vertical structure of gate-oxide-bulk gives the device the Metal-Oxide-Semiconductor (MOS) portion of its name. The substrate serves as the bulk, while the source and drain are formed by doping specific regions of the substrate. The channel is established by applying a voltage between the gate and source terminals ( $V_{GS}$ ), creating a vertical electric field that modulates the conductivity of the channel that facilitates current flow. This modulation is achieved through the formation of a thin inversion layer within the bulk directly underneath the gate oxide. A simple analogy is to view increasing  $V_{GS}$  as a knob that gradually opens the channel in the device, allowing water (the current) to flow between the fittingly named two ends, source and drain (Fig. 2.1b). Applying a voltage between the drain and source terminals ( $V_{DS}$ ) generates a horizontal electric field that attempts to push current through the bulk's channel. Note that the source and drain are identical, and therefore, interchangeable. As suggested by their names, carriers always flow from the source to the drain, and hence, switching the polarity of applied  $V_{DS}$  will switch the source and drain locations, making the MOSFET a symmetric device.

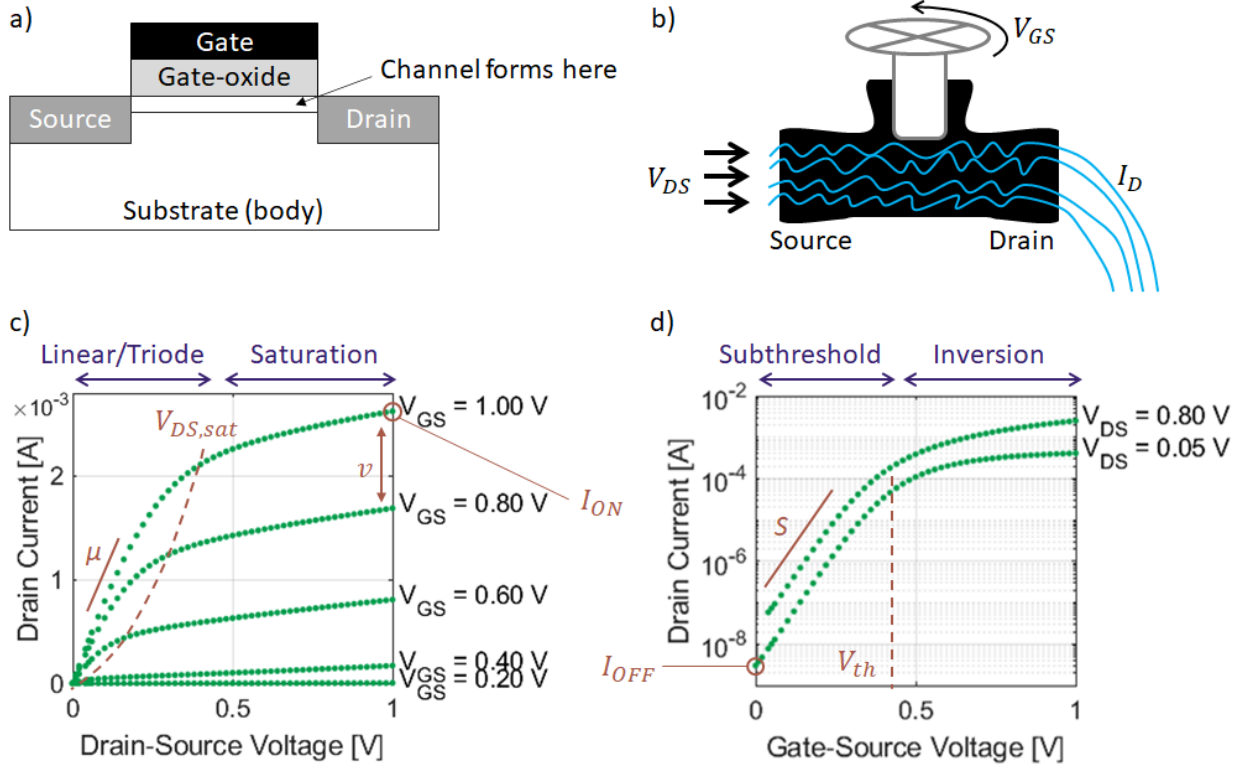


Figure 2.1: (a) Simple illustration of bulk MOSFET cross-section structure. (b) Water analogy to simplify device operation principles. (c) Output characteristics, or  $I_D$ - $V_{DS}$  curves, of a real NMOS device. (d) Transfer characteristics, or  $I_D$ - $V_{GS}$  curves, of the same device.

The device introduced thus far is called an NMOS device, where current is primarily facilitated by electron flow in the channel. Conversely, a p-type MOSFET (PMOS) device employs a p-type channel, where the main charge carriers are holes. It is important to note that current flows in the same direction as holes but in the opposite direction to electrons. As a result, the voltage polarities and current directions in a PMOS are inverted when compared to an NMOS. To prevent confusion, this thesis will mainly focus on the perspective of an NMOS device. Furthermore, in CMOS, the term "complementary" signifies the utilization of both MOSFET types within the same circuit.

To characterise a MOSFET, its performance is examined from two perspectives: the gradual ON-OFF state of the device influenced by  $V_{GS}$  or vertical field (transfer characteristics, Fig. 2.1d), and the current flow in a given state influenced by  $V_{DS}$  or horizontal field (output characteristics, Fig. 2.1c). These two perspectives are typically explored by observing the resultant drain current

( $I_D$ ). While it may be expected that a device with a formed channel can allow any supplied current to pass through, there is a limit to its capacity. Examining the output characteristics in Fig. 2.1c reveals that the current initially increases proportionally with  $V_{DS}$ , but then seems to saturate as it becomes more challenging to push additional current through the channel, even with increased  $V_{DS}$  potential. These two operation regions, linear and saturation, are separated by a critical voltage known as the saturation voltage ( $V_{DSAT}$ ). Note that the saturation level increases when the device is put under higher  $V_{GS}$  potential. As expected, the saturation effect depends on the channel's inversion state and can be modeled by the simplified relationship:  $V_{DSAT} = V_{GS} - V_{th}$ . The simplification arises from defining the threshold voltage ( $V_{th}$ ) as the value at which the channel is abruptly formed when  $V_{GS}$  equals  $V_{th}$ . Predictably, current is cut off when  $V_{GS}$  falls below  $V_{th}$ , as  $V_{DSAT}$  would approach 0 or negative values. It is important to note that channel formation is actually gradual, and this simplified analysis does not encompass the many complex actions that occur in a MOSFET.

From the other perspective, the transfer characteristics display the turn-on behaviour of the MOSFET as  $V_{GS}$  is varied (or the knob is turned). Viewing these characteristics in the linear scale may give the impression of an abrupt turn-on behaviour, but viewing  $I_D$  in the logarithmic scale reveals the rapid but *gradual* channel formation. At a fixed  $V_{DS}$ , it is possible to specify a threshold ( $V_{th}$ ) that separates the graph into two other operation regions. These regions are fittingly named as the subthreshold region (OFF) and the inversion region (ON), and not to be confused by the linear and saturation regions from the  $I_D$ - $V_{DS}$  perspective. In a more realistic graph, the inversion is much more gradual, and these operation regions are seen in a different perspective as weak- and strong-inversion, with  $V_{th}$  lying somewhere in between. Typically, the transfer characteristics include two curves to link the  $I_D$ - $V_{GS}$  perspective to the  $I_D$ - $V_{DS}$  perspective. These two curves provide a window to the linear and saturation operation region, invoked by applying low and high  $V_{DS}$  values, 50 mV and 0.8 V, respectively. (The same technique was utilized previously in the output characteristics, where every curve corresponded to a different  $I_D$ - $V_{DS}$  operation region or  $V_{GS}$  value.)

Both perspectives ( $I_D$ - $V_{DS}$  and  $I_D$ - $V_{GS}$ ) are used to reveal crucial device parameters similar to  $V_{th}$ . Another  $I_D$ - $V_{GS}$  parameter is the subthreshold swing ( $S$ ) which is extracted from the slope of the subthreshold region ( $SS$ ):  $S = 1/SS$ . It characterizes the MOSFET's ability to control the channel conductivity over a wide range of gate voltages. Through this parameter, the leakage current (also known as  $I_{OFF}$ ) can be determined by finding the y-intersect, where low leakage is desired. Similarly, in  $I_D$ - $V_{DS}$ , the slope of the linear region is influenced by the internal resistance of the device ( $slope = 1/(R_{lin})$ ). Note that for a higher fixed  $V_{GS}$ , the implied resistance is lower. This is because the channel is in deeper inversion, making it more conductive and less resistive. At the saturation region, this slope refers to the output-resistance of the device. The saturation region also helps determining the MOSFET's maximum drive current (also referred

to as  $I_{ON}$ ), which directly impacts its ability to drive other circuit components. Together with the leakage current, the  $I_{ON}/I_{OFF}$  ratio provides a clear view of the trade-off between speed and power consumption and can be directly evaluated via the  $S$  performance parameter. Furthermore, the difference in drive-current between two  $V_{GS}$  biases reveals important information about the device regarding its carrier mobility ( $\mu$ ) and velocity ( $v$ ). These two parameters characterize the speed and efficiency of the device:  $\mu$  refers to the ease with which charge carriers move within the channel, influencing the MOSFET's conductivity, while  $v$  represents the speed at which the charge carriers move under the influence of an electric field, determining the MOSFET's switching speed and overall performance. Moreover, a constant difference between two  $V_{GS}$  curves is constant indicates a saturated  $v$ , which occurs in short-channel devices ( $L < \sim 100\text{nm}$ ).

## 2.2 MOSFET at CryoT

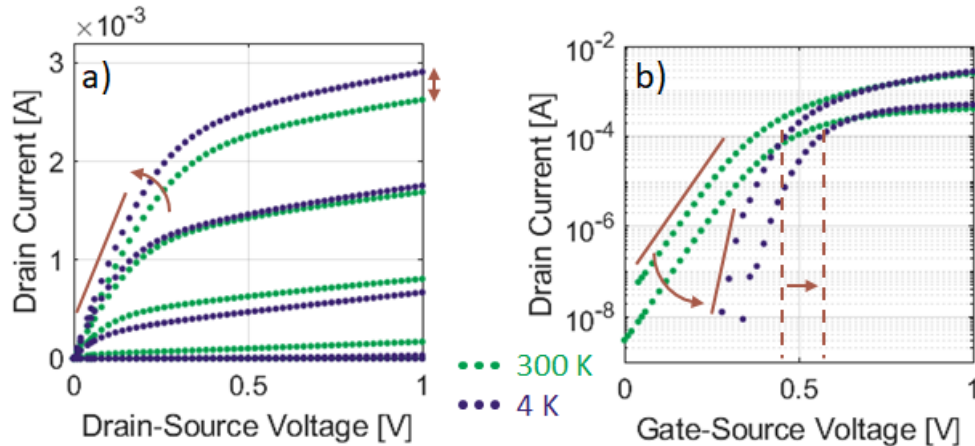


Figure 2.2: Measured  $I$ - $V$  characteristics of a 65-nm bulk NMOS device at RT and cryoT. (a) Output characteristics with  $V_{GS} = [0, 0.2, 0.4, 0.6, 0.8, 1]\text{V}$ . (b) Transfer characteristics with  $V_{DS} = [0.05, 0.8]\text{V}$ .

MOSFETs are manufactured to be utilized in everyday electronics such as cell phones, computers, cars, and even super computers. They are meticulously engineered to not only withstand a wide range of operating temperatures but also meet stringent performance criteria. The typical temperature range spans from  $-55^\circ\text{C}$  to  $200^\circ\text{C}$  (or 218 K to 473 K), encompassing a broad spectrum but still far from extremely low temperatures. Nevertheless, commercial MOSFETs are, in fact, capable of operating at cryoT, albeit with altered performance as shown in Fig. 2.2.

Indeed, this is possible because the fundamental principle of operation and first-order physics are still the same at cryoT, which is why it is possible to adopt the current Si infrastructure and design principles.

As temperature is reduced, the performance parameters of a MOSFET device are altered in various ways. Fig. 2.2 presents these variations in a 65-nm bulk NMOS device when the operating temperature of is decreased form RT to cryoT. The increase in  $V_{th}$  seen in Fig. 2.2b suggests the need for higher operating voltages (and consuming more power), but the sharper subthreshold slope at cryoT indicates a higher  $I_{ON}/I_{OFF}$  ratio, meaning that the device exhibits low power leakage capabilities, and is more efficient overall. Moreover, despite the strong inverse relation between the drive current and  $V_{th}$ , Fig. 2.2a shows notable increase in the former at cryoT without the need to operate at higher voltages. This is important because a certain level of drive current is required for high performance operation, as it directly influences the technology's speed. Achieving high  $I_{ON}$  and low  $I_{OFF}$  simultaneously by reducing the temperature is very attractive for circuit design. This is especially the case when the operating voltage (a key design parameter) does not need to change, making it easier to convert RT designs in cryoT. Furthermore, note that the high cryoT current level cannot be explained by the reduced internal resistance alone. This behaviour indicates a notable increase in transport parameters, namely  $\mu$  and  $v$ , which can be correlated to the inverse relation between phonon scattering and temperature. At first glance, some of these results seem to corroborate classical theories predictions, but closer inspection reveals the necessity to re-examine widely accepted models.

One of the critical performance parameters is the subthreshold swing, given that lower  $S$  values allow for more performance at lower leakage. This significance becomes particularly pronounced in cryo-electronics, where escalated power consumption results in elevated temperatures, consequently disrupting qubit operation. Modern FET technologies continuously strive to maintain low subthreshold swing or even improve it beyond its theoretical limit of 60 mV at RT.

Fortunately,  $S$  improves at lower temperatures, making it more attractive to operate in the cryogenic regime. Therefore, it is important to accurately model  $S(T)$  and understand its deviation from the classical model that paints a linear relation with temperature:

$$S = \frac{\partial V_G}{\partial \log(I_D)} = m \ln(10) \phi_T \quad (2.1a)$$

$$m = \frac{\partial V_{GS}}{\partial \phi_s} \quad (2.1b)$$

where the subthreshold factor  $m$  is a dimensionless constant that describes the effect of  $V_{GS}$  on the surface potential  $\phi_s$ . The thermal voltage  $\phi_T = k_B T / q$  is related to  $k_B$ : the Boltzmann constant and  $q$ : the elementary charge. As its name suggests,  $\phi_T$  is directly proportional to  $T$ , while the

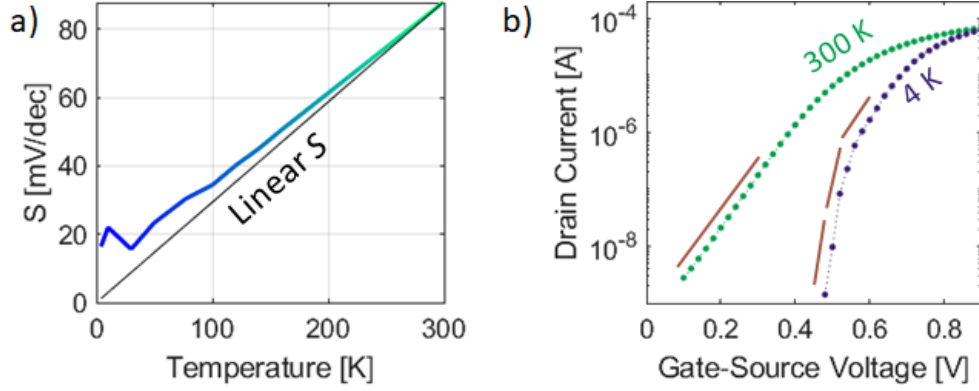


Figure 2.3: (a) Measured subthreshold swing vs. temperature (colored solid line) compared against the classical model with  $m = 1.48$  (solid black line). (b) Measured transfer characteristics at RT and cryoT, showing the variable  $m$  effect. Measurements were conducted on a 65-nm bulk NMOS sample with  $V_{DS} = 50mV$ .

other values are constants and unaffected by  $T$  ( $k_B = 1.38 \times 10^{-23} J/K$  and  $q = 1.6 \times 10^{-19} C$ ). Therefore, it is expected for  $S$  to improve by 75 times when the transistor is cooled down from  $T = 300K$  to  $T = 4K$ , but the results shown in Fig. 2.2b indicate an improvement of only  $\sim 7mV/dec$ . In fact, Fig. 2.3a reveals a subthreshold saturation effect in bulk CMOS at deep cryoT. This effect is also observed in more advanced nodes such as Fully-Depleted SOI (FDSOI) [33]. Moreover, evidence of a non-constant subthreshold factor  $m$  at cryoT have been observed in cryogenic I-V measurements [28]. A constant subthreshold factor leads to a constant  $S$  as seen in the RT transfer curve (Fig. 2.3b). However, variability in  $S$  can be observed well below  $V_{th}$  in the cryoT  $I_D$ - $V_{GS}$  curve in the same figure, suggesting a gate-voltage dependence.

Another example of deviation from the expected behaviour is the threshold voltage response to decreasing temperature.  $V_{th}$  is commonly modeled to respond linearly with temperature: [34]:

$$V_{th}(T) = V_{tho} + \alpha(T - T_o) \quad (2.2)$$

where  $T_o$  is a reference temperature where the reference threshold voltage is defined ( $V_{tho} = V_{th}(T_o)$ ), and  $\alpha$  is a temperature coefficient, estimated to be  $\sim 2mV/K$  for Si devices. Measurement results of  $V_{th}$  down to cryoT are displayed in Fig. 2.4 where the linear relationship is apparent for temperatures above  $\sim 150K$ , but the trend drastically changes in the cryogenic regime. Indeed, the simple model eq. 2.2 relies on multiple approximations that discount physical terms negligible at a higher temperatures, which can account for the observed discrepancy. CryoMOSFET models must revisit the full physical definition while carefully handling easy-to-make approximations, such as complete ionization and analytical statistics (e.g. Maxwell-

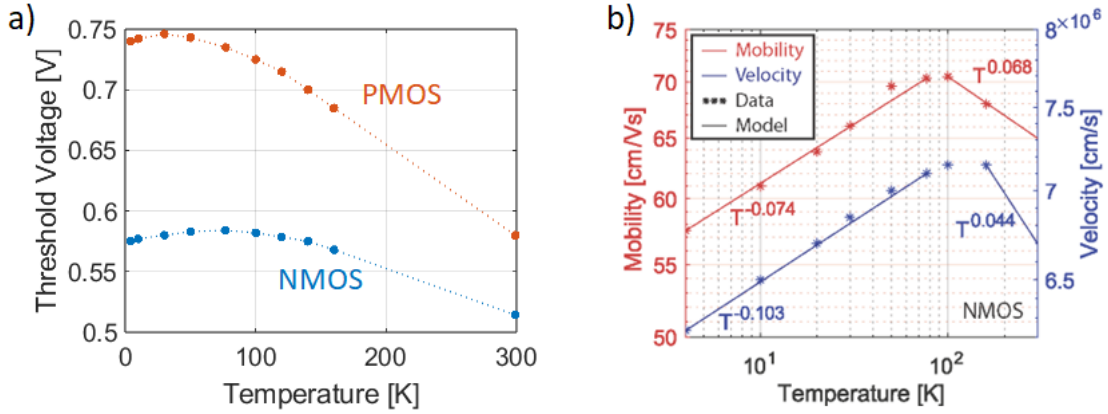


Figure 2.4: Measured data of (a) threshold voltage and (b) transport parameters vs. temperature in 65-nm bulk devices.

Boltzmann vs. Fermi-Dirac statistics).

Regarding transport parameters,  $\mu$  is generally expected to improve with reduced temperature due to reduced phonon scattering [35]. Measurement results (Fig. 2.4) reveal that this seems to be generally the case until a certain temperature, and then the trend starts changing. This can be attributed to the existence of other scattering mechanisms that have different relations to temperature. For example, Coulomb scattering is inversely proportional to temperature (making its mobility component proportional to temperature) [35]. Moreover, the effective and ballistic mobility components are directly proportional to temperature through the thermal velocity,  $v_T$ , [36]. Therefore,  $\mu$  is expected to be a strong function of the manufacturing process (e.g. strain) and difficult to model. Similarly, carrier velocity displays similar inconsistencies which can be affected by the similar processes as mobility.

Finally, an unexpected effect that appears in cryoT is the kink effect 2.5a. As illustrated in Fig. 2.5b, the effect originates from increased bulk resistance at cryoT, which induces higher bulk potential, reducing  $V_{th}$  through the body effect, and increasing  $I_D$ . The hysteresis shown in Fig. 2.5c is byproduct of this effect where the forward sweep captures the true  $V_{th}$  before the kink effect activates, while the backwards sweep captures the reduced  $V_{th}$ . However, the kink effect requires high drain voltages to activate  $V_{DS} \geq 1.5V$  which are much higher than the operating voltages of process nodes relevant to this research ( $> 350\text{-nm}$ ) [21]. Hence, this effect is not included in the model as it does not occur within the possible operation range.

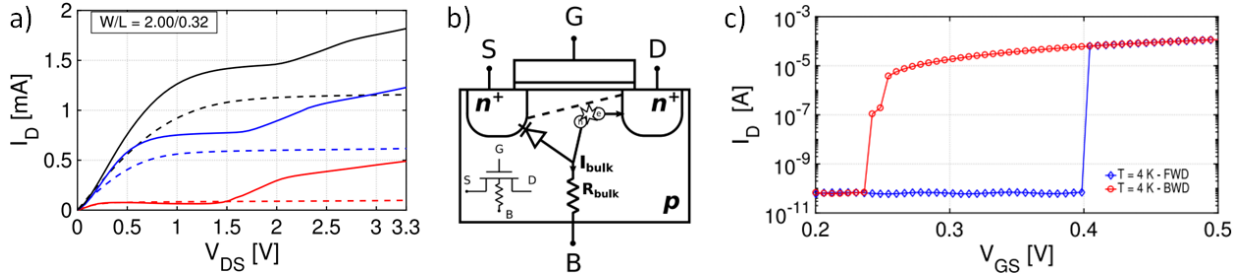


Figure 2.5: (a) Output characteristics of a 160-nm bulk NMOS device at 300 K (dashed) and 4 K (solid). (b) NMOS cross section representation of the origin of the kink effect at cryoT. (c) Transfer characteristics 4 K showing the hysteresis effect. These figures are originally from the work published in [21].

## 2.3 Compact Models

With the knowledge that the physical operation principles of MOSFETs are maintained at low temperatures, it is reasonable to pursue the utilization of this technology in cryo-electronics. Capitalizing on the existing semiconductor infrastructure and circuit design principles that have been continuously developed for decades presents a compelling opportunity for the development of innovative cryogenic IC solutions. In the IC design process, circuit simulations are carried out by circuit simulation engines known as Simulation Program with Integrated Circuit Emphasis (SPICE). SPICE simulators employ numerical methods predict the circuit's overall behavior. These simulators analyze the circuit's response to different stimuli, allowing designers to optimize circuit performance and verify its functionality before fabrication. In order to be represented in the circuit, each electrical component is expressed with a set of closed-form equations, known as compact models, that describe its electrical behaviour (Fig. 2.6).

Compact models are a vital component in the IC design process, and are an indispensable tool for circuit designers in today's fast-paced and ever-changing technology landscape. They serve as a bridge between the complex physical behaviour of an individual device and circuit simulators that manage a large number of devices integrated together in a circuit. This is accomplished by providing a simplified representation of the component's electrical characteristics, allowing for efficient simulation of large-scale circuits.

Compact models must meet several requirements to ensure that designers can confidently and efficiently simulate large circuits. They must be flexible enough to support the wide-range of design parameters expected in large-scale ICs (e.g. device width and length) as well as process variation used by circuit designers to gauge circuit reliability against process uncertainties. In order for designers to have confidence in their simulations, compact models must also be accurate

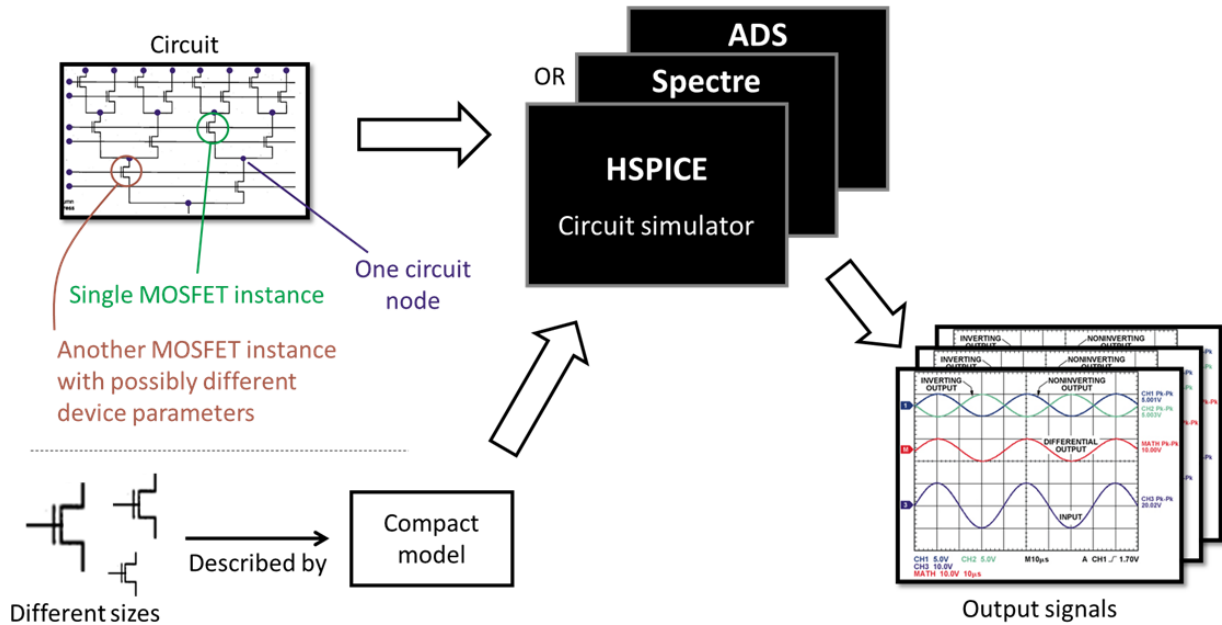


Figure 2.6: A simplified diagram showing the how compact model fits within circuit simulation and IC design flow. Circuits are made of many circuit components including MOSFETs and possibly other elements like resistors and capacitors. Circuit simulators have to solve for all nodal voltages (and the branch currents between them) for large-scale complex circuits. For context, the circuit shown here is considered small-scale and simple. Compact models must be able to represent all transistors present in the circuit, which may have different parameters (represented here with size only). Compact models must also not target one SPICE simulator, and be compatible with a wide range of circuit simulators.

over a wide range of operating conditions, while still being computationally efficient to allow for large numbers of devices in complex systems. Moreover, compact models must be modular to support different simulators used in the industry, like Spectre and Simulation Program with Integrated Circuit Emphasis (SPICE). They must not rely on a specific simulator's features and be able to produce the same results in different simulators, as well as ensure smooth compatibility with older and newer versions of the same simulator engine.

There are several industry-standard compact models, pioneered by the BSIM models [37], that are used to simulate the electrical performance of state-of-the-art devices and integrated circuits. Semiconductor foundries, like Taiwan Semiconductor Manufacturing Company (TSMC) and GlobalFoundries (GF), package these compact models as part of the PDKs of each process node. However, these models are typically developed for device operation at RT and may

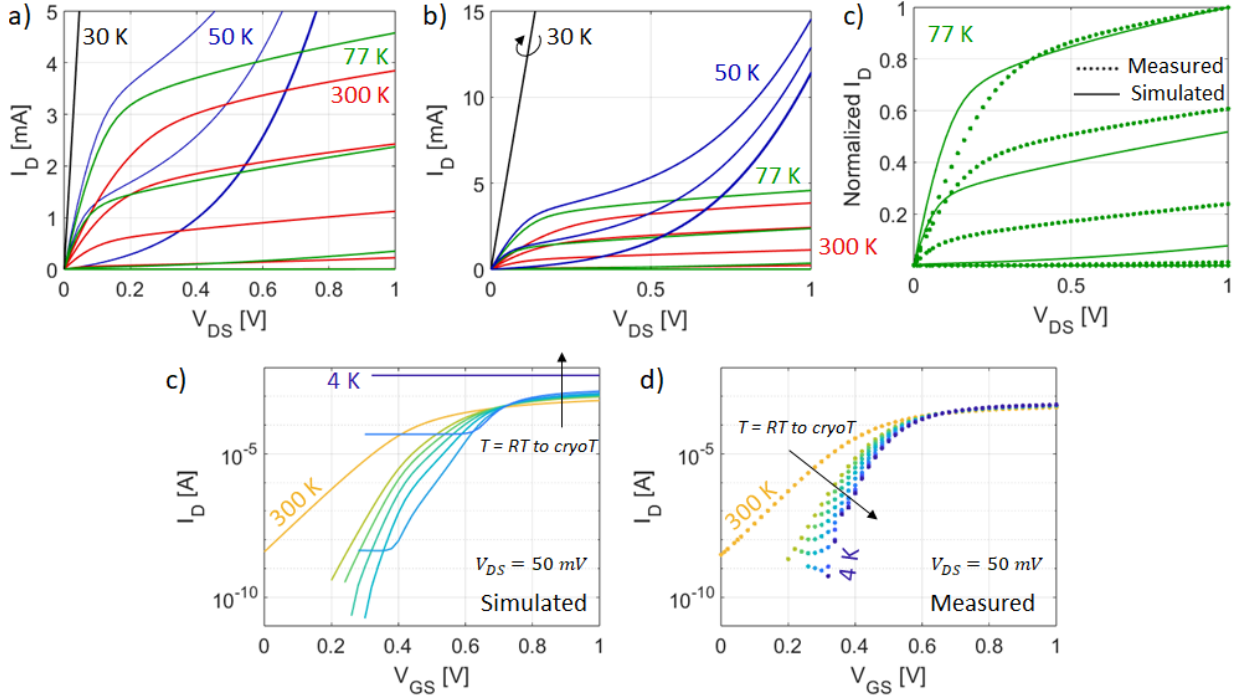


Figure 2.7: I-V curves of a 65-nm NMOS device with  $W/L : 5\mu\text{m}/65\text{nm}$  from the TSMC N65 GP process node. Simulation using the PDK (solid lines) and measurements (lines with markers) are depicted. (a)-(b) Display the clear inability of the PDK model to describe the device behaviour in deep cryoT ( $< 77\text{ K}$ ). Zoomed in (a) and zoomed out (b) perspectives are shown. Since the 77 K simulation display potentially plausible results, it is compared in (c) against experimental results of the same device at the same temperature. Normalized current is obtained using:  $I_{D,normalized} = I_D / \max(I_D)$ . (d)-(e) Highlight the discrepancy between simulation and measurement results through the linear transfer characteristics at multiple temperature points.

be applicable to devices operating down to 150 K. However, they are expected to fail at deep cryoT. The cryogenic capability of the model included in the TSMC general purpose 65-nm bulk technology (TSMC N65GP) is tested. Simulation results generated by popular circuit design platform, Cadence, are shown in Fig. 2.7. While the 77K simulation appear to be plausible, unrealistic results started forming at lower temperatures (Fig. 2.7a), indicating the model's shortcoming at deep cryogenics. Moreover, experimental measurements of the same device at 77 K are used to validate the simulation accuracy at this temperature point. Since the simulations used do not include layout effects (known as pre-layout simulations), the currents are normalized in Fig. 2.7b to minimize the impact of layout, pad, wire bonding, and various measurement setup

parasitic. The vast difference between simulated and measured results is a clear indication of the model’s failure to accurately describe the device at 77 K. Moreover, the exaggerated  $V_{th}$  shift and subthreshold performance inaccuracies are apparent from Figs. 2.7c and d. These results present a clear need to develop accurate compact models in order for cryo-electronics to truly harness the full potential of CMOS technologies.

## 2.4 Progress in Cryogenic Compact Models

As demonstrated in section 2.3, commercial compact models (CMs) are incapable of predicting device performance at cryogenic temperatures, rendering cryoT circuit simulations invalid. This forced early low-temperature circuit development to manually adjust the foundry-provided model parameters (typically the BSIM3 model [38]) to fit cryoT measurements to the best of the model’s capabilities ([9, 31, 39, 40]). These adjustments were often done according to output characteristics alone, and avoided the kink in old technologies ( $> 300$ -nm process nodes) by restricting operation to low voltages.

Nevertheless, there has been notable effort to bring industry-standard compact models close to the cryogenic regime. Several standard models received this treatment (MOS11[41], PSP[42], BSIM [37], EKV[43], MVS[44]) either by simply adjusting model parameters, or by modifying device equations to include temperature dependence and additional cryoT effects as summarized in Table 2.1. There were also attempts to develop original proprietary models such as the XMOS model [30]. Moreover, some effort was made to develop methods devised specifically for cryogenic extraction, and tailored towards popular models like the EKV and BSIM [45]. It is found that the EKV2.6 model is more capable of matching cryoT I-V characteristics than the BSIM3v3 model. This finding is corroborated by [22, 46], where the simplified EKV model is declared able to accurately predict the cryogenic behavior through parameter adjustment alone. W. Chakraborty et al. was able to adjust the MVS model parameters to fit multiple temperature points, down to 6 K, though no comparison with other models was made. This approach, however, requires treating each temperature measurement as an individual device with its own parameter set. This tedious approach is equivalent to fitting the same number of devices even if they were from different process nodes. Moreover, fitting results show a clear need for updating device equations to address cryoT exclusive effects, like  $S$ -saturation [47] and  $V_{GS}$ -dependence [28].

To this end, several models were modified to include some cryoT effects. R. M. Incandela et al. included the kink effect as a non-linear resistor connected in series with the bulk (first explained in [48]), and combated the cryoT subthreshold swing discrepancy by adjusting the factor,  $n$ , using a look-up table [21, 49]. These modifications were made to the MOS11 and

PSP models, with the PSP prevailing in accuracy. The industry-prominent model group, BSIM, was also modified by several groups to bring it closer to the cryogenic regime. Temperature coefficients were added to the BSIM3 model equations [50–52], the effective  $V_{th}$  equations was modified in BSIM4 [53], and an approach with emphasis on effective mobility ( $\mu_{eff}$ ) was taken to modify the BSIM-CMG model [54]. Moreover, a BSIM-based model that included many cryoT effects (such as  $S$ -saturation) was developed by G. Pahwa [55]. It can be noticed that the kink effect received little to no attention since scaled down process nodes operate at much lower voltages, and therefore, are practically immune to this effect. The EKV3 model was also modified to include temperature dependence [56], but it lacks the inclusion of cryoT exclusive effects.

Despite the extensive efforts to produce a cryoT CM, complex cryogenic effects require rigorous investigation of the underlying physics before incorporating them into CMs. A. Beckers et al. published a physical review of the MOSFET device physics in cryogenics, and provided core expressions for the development of physically accurate cryoT CMs [24]. Further, several groups studied the cryogenic influence on various aspects of device performance, including:  $V_{th}$  [57–59],  $S$  [25, 28, 47, 60–62], and  $\mu$  [15, 29, 32, 35, 63, 64], where they characterized device performance, explored physical explanations, and even created physical models. J. Hoff et al, explored device reliability through stress-testing in cryogenic environments [65].

Table 2.1: Survey of the state-of-the-art in cryoMOSFET models.

| Ref  | Compact Model    | Focus                                 | Circuit Demo | Min. Temp. | Model Edit Type                |
|------|------------------|---------------------------------------|--------------|------------|--------------------------------|
| [9]  | MOS11            | -                                     | LNA & OSC    | 4 K        | Par. adj.                      |
| [39] | BSIM3v3          | -                                     | ADC          | 4.2 K      | Par. adj.                      |
| [31] | BSIM3.3          | -                                     | V-reference  | 77 K       | Par. adj.                      |
| [40] | Not-mentioned    | -                                     | ADC          | 4.4 K      | Par. adj.                      |
| [30] | XMOS             | Physics-based                         | -            | 8 K        | Original CM                    |
| [45] | BSIM3v3 & EKV2.6 | -                                     | -            | 77 K       | Par. extraction                |
| [22] | Simplified EKV   | $S$ vs. $T$ table                     | -            | 4.2 K      | Par. adj.                      |
| [46] | Simplified EKV   | cryoT electrostatics<br>$S$ -offset   | -            | 4.2 K      | Par. adj.                      |
| [28] | MVS              | cryoT $S$<br>$S$ vs. $V_{GS}$         | -            | 4.2 K      | Additional eqs.<br>& par. adj. |
| [49] | MOS11<br>BSIM3v3 | Kink effect eqs.<br>$S$ vs. $T$ table | RO           | 100 mK     | Additional eqs.<br>& par. adj. |

| Ref      | Compact Model | Focus                                              | Circuit Demo | Min. Temp. | Model Edit Type             |
|----------|---------------|----------------------------------------------------|--------------|------------|-----------------------------|
| [21]     | MOS11         | Kink effect eqs.                                   | RO & LNA     | 100 mK     | Additional eqs. & par. adj. |
|          | BSIM3v3       | $S$ vs. $T$ table                                  |              |            |                             |
| [50]     | BSIM3v3       | -                                                  | -            | 4.2 K      | Par. adj.                   |
| [52]     | -             | Incomplete ionization & $\mu$                      | -            | 30 K       | Device simulator            |
| [51]     | BSIM3         | $S$ , $V_{th}$ , and $\mu$ drift-diffusion based   | -            | 20 K       | Additional eqs.             |
| [53]     | BSIM4         | $S$ , $V_{th}$ , and $\mu$ kink effect             | -            | 4 K        | Additional eqs.             |
| [54]     | BSIM-CMG      | $S$ -saturation, $V_{th}$ , $\mu$                  | -            | 10 K       | Modified eqs.               |
| [55]     | BSIM-based    | $S$ -saturation, $V_{th}$ , $\mu$ $S$ vs. $V_{GS}$ | -            | 4.2 K      | Modified eqs.               |
| [56]     | EKV3          | Model evaluation                                   | -            | 100 mK     | Modified eqs.               |
| [26]     | MVS           | Trend explanation                                  | -            | 6 K        | Par. adj.                   |
| [24]     | -             | Physical model for cryoT devices                   | -            | 4 K        | Original model (not CM)     |
| [57]     | -             | $V_{th}$ physical characterization                 | -            | 173 K      | -                           |
| [58]     | -             | $V_{th}$ physical model                            | -            | 6 K        | -                           |
| [59]     | -             | $V_{th}$ physical model                            | -            | 4.2 K      | -                           |
| [25, 60] | -             | General $S$ cryoT models                           | -            | 4.2 K      | Original models (not CMs)   |
| [62]     | -             | $S$ -saturation                                    | -            | 2 K        | -                           |
| [47, 66] | -             | $S$ -saturation                                    | -            | 4.2 K      | -                           |
| [61]     | -             | $S$ -inflection                                    | -            | 4.2 K      | -                           |
| [15]     | -             | $S$ & $\mu$                                        | -            | 20 mK      | -                           |
| [29]     | -             | $V_{th}$ , $\mu$ , $D_{it}$                        | -            | 100 mK     | -                           |
| [35]     | -             | $V_{th}$ , $\mu$                                   | -            | 10 K       | -                           |
| [64]     | -             | Neutral defects                                    | -            | 50 K       | -                           |
| [32, 63] | -             | $V_{th}$ , $\mu$ , $S$ & Neutral defects           | -            | 77 K       | -                           |
| [65]     | -             | MOSFET lifetime in cryoT                           | -            | 77 K       | -                           |

All these works underscore the need for deeper comprehension of the intricate physics behind cryoT phenomena and, concomitantly, the formulation of a unified compact model that encapsulates all cryoT phenomena at the same time. This research will explore cryoT phenomena, and incorporate the findings into one model, the MIT Virtual-Source (MVS) model (chapters 3-4). MVS is a flexible CM suitable for prototyping, technology benchmarking that, and first-pass circuit design simulations. To the best of our discernment, no prior investigations have attempted to update the temperature range of this model yet, making this study unique in that regard. The CM development must also consider practicality and ease of adoption. For this reason, extraction tools are developed alongside the model to lower the barrier-to-entry into cryo-electronics design (chapter 5). Moreover, validation through circuit simulation is vital to test and showcase the CM's capabilities. After going through the extraction process, circuit simulation of a Demultiplexer circuit (Demux) circuit common in Quantum computing applications [67] is demonstrated.

## Chapter 3

# Cryogenic Measurements

The silicon technologies utilized in mainstream electronics (such as smartphones and computers) are developed and rigorously tested at temperatures suitable for their intended operation. Typically, these temperatures span a range of  $-50^{\circ}\text{C}$  to  $200^{\circ}\text{C}$  (or 223.15 K to 473.15 K), which falls considerably short of the cryoT regime (e.g. liquid nitrogen's temperature is 77 K or  $-196.15^{\circ}\text{C}$ ). Despite that, these technologies are still capable of operating in extreme temperatures, even below 4.2 K (liquid helium temperature) [29]. Nevertheless, the exact physics at cryoT are not fully understood yet, and the underlying device models lack consideration for such temperatures and are unlikely to account for any exclusive cryoT effect. Consequently, acquiring raw experimental data from real-world observations becomes imperative, allowing for exploration of this uncharted territory without being confined by theoretical biases. By obtaining unbiased experimental data, it becomes possible to unravel the intricacies of device behavior to validate theoretical predictions and uncover any phenomena exclusive to cryoT. These data-driven insights serve as the foundation for enhancing our understanding of silicon technologies in cryogenic regimes.

At cryogenic temperatures, even the simplest measurements quickly become complex and deeply involved. For instance, submerging a sample (the DUT) in liquid helium necessitates stringent safety measures. Helium is a scarce and very expensive resource on Earth, and it expands 750 times when transitioning from liquid to gas [68]. A casual fast immersion would not only boil off large amounts of precious helium, but in extreme cases, the expanded volume of the newly formed gas produces enormous pressure, enough to catastrophically explode the containing dewar. Owing to its scarcity, suitable infrastructure is in place to capture and re-liquefy the helium gas inevitably produced, and therefore, great care must be taken while loading and unloading samples to avoid contamination.

A major electrical consideration is the dipping probe, which becomes an integral part of the

measurement system, necessitating thorough understanding of its properties, as explored in section 3.2. Furthermore, the expected subthreshold current levels at cryoT are orders of magnitude smaller than at RT, which are difficult to detect and isolate, requiring specialized research-grade measurement equipment. Fortunately, this research project benefits from the availability of a dedicated cryogenic laboratory equipped with research-grade instruments specifically designed for quantum computing research. This laboratory, situated at the RAC on our university campus, is led by Professor Jonathan Baugh, a professor at IQC and the University of Waterloo’s Department of Chemistry. This research endeavor is a collaboration with prof. Baugh’s research team who offered their invaluable insight and expertise to realize the measurements discussed in this chapter.

### 3.1 Measurement Samples

The samples analyzed in this study are based on multiple bulk CMOS process nodes ranging from 65-nm to 20-nm. As discussed in chapter 2, such mature technologies are attractive for quantum computing cryo-electronics. They can provide adequate performance and compact chip area at a lower cost and higher yield with flexible layout rules and better access to talent compared with more advanced technologies such as FinFET.

Table 3.1: Summary of measured samples.

| Sample | Process Node | Name   | Type**    | Width<br>( $\mu m$ ) | Mask Gate<br>Length (nm) | Cox<br>( $mF/m^2$ ) |
|--------|--------------|--------|-----------|----------------------|--------------------------|---------------------|
| TSM65  | TSMC N65     | TSM65n | NMOS-HVT  | 4.8                  | 65                       | 12.68*              |
|        |              | TSM65p | PMOS-SVT  | 7.2                  |                          | 15.01*              |
| TSM28  | TSMC N28     | TSM28n | NMOS-SVT  | 1.2                  | 30                       | 26.56*              |
|        |              | TSM28p | PMOS-SVT  | 1.2                  |                          | 26.56*              |
| TSM20  | TSMC N20     | TSM20n | NMOS-ULVT | 0.72                 | 20                       | 31.39*              |
|        |              | TSM20p | PMOS-ULVT | 0.84                 |                          | 28.78*              |
| CMC65  | TSMC N65 G   | CMC65n | NMOS-SVT  | 5                    | 65                       | 12.68               |
|        |              | CMC65p | PMOS-SVT  | 5                    |                          | 15.01               |

\*Estimated value due to lack of specific information.

\*\*The letter proceeding ‘VT’ indicates the threshold voltage flavor (or level) where: S, H, UL refer to standard, high, and ultra low. Typically, the order of the threshold voltage in the same process node is: ULVT < SVT < HVT. These flavors are not comparable across process nodes and highly technology dependant.

The structures needed for this study are n- and p-type transistors with different gate lengths and made in different technologies. There are two main sample groups involved in this research as listed in Table 3.1. The first group is sample chips sourced from TSMC (dubbed TSM) which are made in 65-nm, 28-nm, and 20-nm process nodes. These test chips are made for automated industry measurements, introducing further challenges when measured manually for measurements they are not designed for (details in section 3.2). To alleviate these challenges, a second group is specially designed for cryogenic measurements and is produced through CMC Microsystems services (dubbed CMC). Although the CMC65 chip uses the TSMC’s general purpose 65nm process [69], it is not equivalent to the TSM65 chips. While little is known about the TSM samples due to protection of trade secrets, measurements results from both groups are different. There are several potential causes for these differences, such as: layout effects, the specific metalization stack used in the fabrication run, maturity of the process (process revision), pads used, different process node flavors (e.g. GP vs LP process technologies), etc. It is possible that the TSM samples are from early or pre-production samples fabricated during the technology development process (around 2005 or before), while the CMC samples are made using the perfected fabrication process after decades of improvements (in 2020).

Overall results from TSM samples suffered from reliability issues where samples often failed during cryoT measurements or even while conducting routine RT measurements beforehand. These failures could be caused during a wide range of stages, including sample preparation, sample handling, and on-chip sample protection. The lack of information about the specifics of the test structure design (such as layout) in the TSM samples makes it difficult to identify failure sources and improve the measurement yield. Thus, several strategies are employed to attempt to overcome these issues, such as: designing a test chip specifically for cryoT measurements (section 3.1), utilizing creative wire-bonding techniques, and using more advanced cryogenic systems (section 3.2).

### 3.1.1 CMC65 Chip Design

The new chip is developed to resolve the following challenges:

1. Determining whether electrostatic discharge (ESD) protection is implemented or not, and its effect, if any, on the results.
2. The input/output (I/O) pad placement and the tight pitch utilized in the TSM samples. This is related to wire-bonding yield and complexity.
3. Unknown structure layout, including well layout and device isolation, guard-ring layout, shared connections between devices (like gates and body).

4. The lack of an on-chip temperature sensing solution.
5. The lack of general knowledge about the technology used.

The CMC65 chip is designed using the PDK for TSMC's 65-nm general purpose (GP) process provided by CMC Microsystems [69]. It is made to produce known DUT structures and isolate failure points. Additional test structures are added to help isolate failure points, such as structures with and without protection circuits, and an on-chip temperature sensor. The chip also contains cryogenic test circuits for other aspects of the cryogenic CMOS project that are beyond the scope of this thesis. Nevertheless, designing a chip comes with a multitude of practical challenges that must be addressed.

### ESD Protection

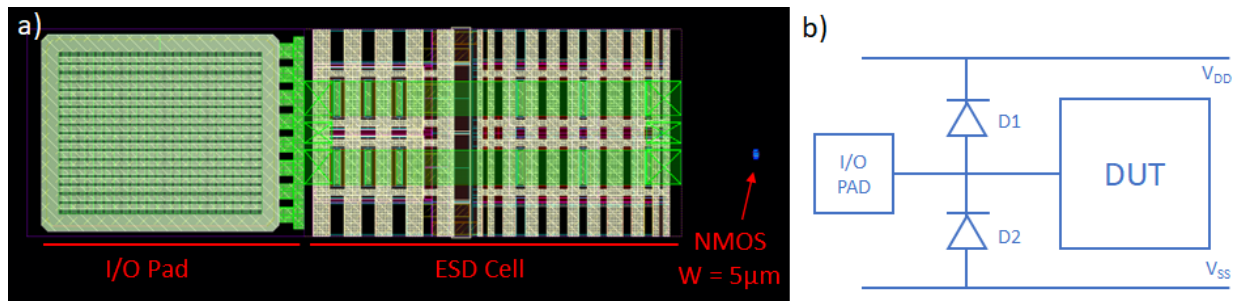


Figure 3.1: (a) Area comparison between I/O pad, ESD Cell, and a typical DUT. (b) An example of an ESD protection circuit.

A possible reason behind the TSM samples' sensitivity to excessive handling is ESD, a major concern in CMOS chip design in general. ESD is a sudden flow of electrical charge between two objects at different potentials. Such discharge may occur during chip handling, wire bonding, or other activities, and it can generate significant voltage spikes that have the potential to damage the internal test structures. To prevent this, I/O pads are typically not directly connected to the internal devices and circuits on the chip. Instead, ESD protection circuits, like the one shown in Fig. 3.1a, are employed in between. Any voltage above  $V_{DD}$  or below  $V_{SS}$  is rapidly discharged through D1 or D2, respectively, before it propagates further into the chip and damages its components. ESD protection circuits are provided in the PDK (referred to as ESD cells) and it is unknown if the TSM samples include them or not.

ESD cells come with a disadvantage related to area as they are made considerably large for optimal protection. For context, an ESD cell has an area of about 1.5 times that of an I/O pad and

more than 1300 times larger than a CMC65n device. Moreover, ESD cells are additional devices that are considerably larger than the MOSFET being tested, and may affect the results. This is particularly concerning at cryogenic temperatures, where the behavior of the device is not yet fully understood essentially leading to a two-variable one-equation problem. For these reasons, two sets of NMOS and PMOS devices are included in the chip; one set with ESD protection and another without, as marked in Fig. 3.2. The unprotected set is intended as an experiment to ascertain if ESD protection is indeed needed and a backup option if ESD cells interference is significant. The device collection includes multi-finger NMOS and PMOS devices with a base width of  $5 \times 1 \mu m = 5 \mu m$  and a layout gate length of  $65 nm$ . An array of different device lengths and widths is also included, but only in the ESD protected set due to area limitations.

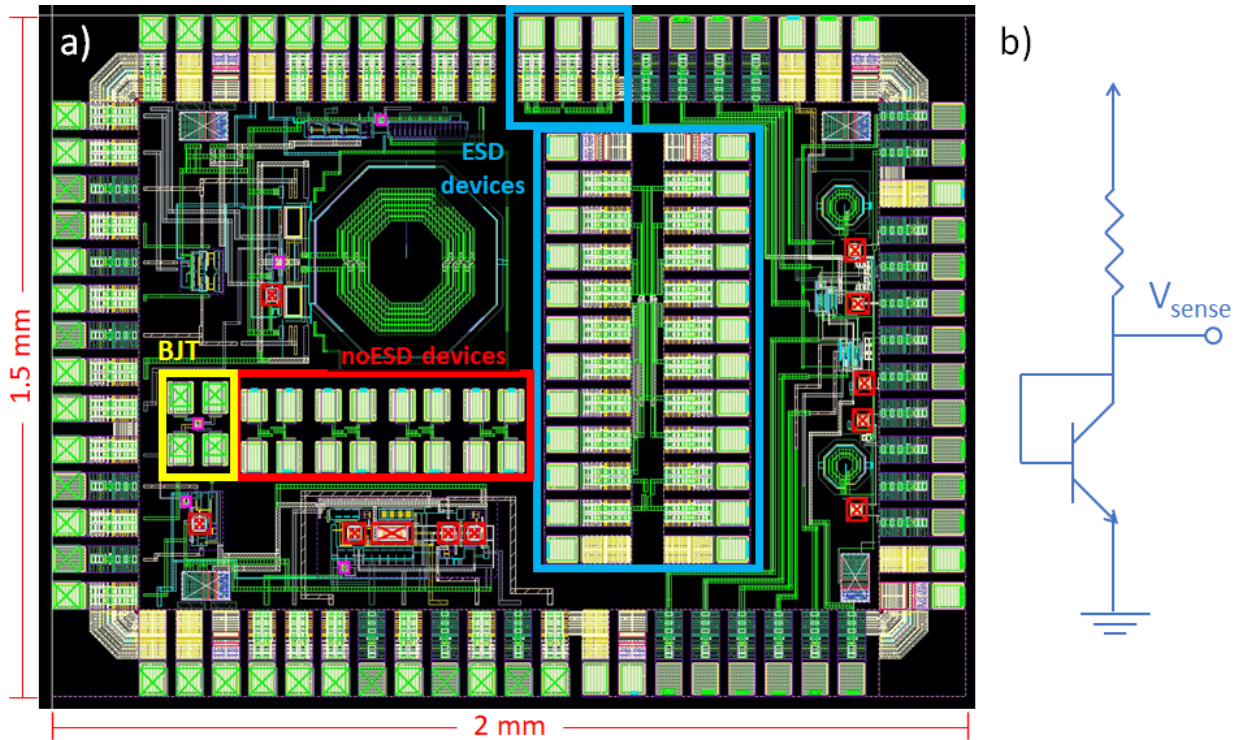


Figure 3.2: (a) CMC chip layout with an area of  $3 mm^2$  where ESD cells and I/O pads occupy 20% and 16% of total chip area, respectively, excluding spacing. (b) A simple temperature sensor circuit made from a diode connected npn Bipolar Junction Transistor (BJT).

## I/O Pad Layout

The chip shown in Fig. 3.2 is surrounded by an I/O pad ring that connects the outside world to the internal structures (stand-alone devices and circuits) with ESD cells in-between. Typically, the pads are placed as close as possible to each other for area efficiency since chip packaging is usually done via high precision automation. Nevertheless, additional spacing is inserted to increase the minimum pad pitch to  $110\mu m$  (or minimum inter-pad spacing of  $53\mu m$ ) to better support manual wire-bonding during sample preparation. The extra spacing can help with wire-bonding yield and simplifies the wire angles, allowing for more pads to be wired simultaneously when needed for complex measurements. Due to area limitations (or more specifically, limited perimeter), part of the I/O-ESD pad structure is moved internally. Additionally, a set of I/O pads that lack ESD protection are also placed internally. These pads connect directly to the "noESD" device set meant to test the need for such protection given our sample preparation and measurement process.

## Device Layout

Each device structure is fully isolated from other devices despite requiring more I/O pads. This is done by separating wells, guard rings and similar signals like gate and body, as illustrated by Fig. 3.3. Moreover, the interconnect layout introduces additional parasitics that must be accounted for in order to achieve desired results. These parasitics can be extracted by the EDA tool and added to circuit simulation to produce what is known as the "post-layout" simulation. Fig. 3.4a shows the difference between pre- (schematic only) and post-layout output characteristics of the same transistor. The post-layout simulation shows slightly lower current levels owing to additional interconnect resistance. Fig. 3.4b shows the importance of optimized layout to minimize layout parasitics by comparing simulation results of the same transistor implemented using two different layouts. Using post-layout simulation brings simulation results closer to post-silicon performance, and improves the yield.

## On-Chip Temperature Sensor

As outlined in greater detail in section 3.2, the sample will not be in direct contact with the cryogenic liquid, but instead will be housed in a vacuum can, thereby significantly reducing heat transfer and slowing down the cooling process. While the temperature inside the cryostat is known, it is important to create a baseline of the time it takes the chip to change temperatures and achieve thermal equilibrium. The sensors on the temperature-controlled cryostat cannot be relied on because they are placed outside the vacuum can and cannot provide any insight on the

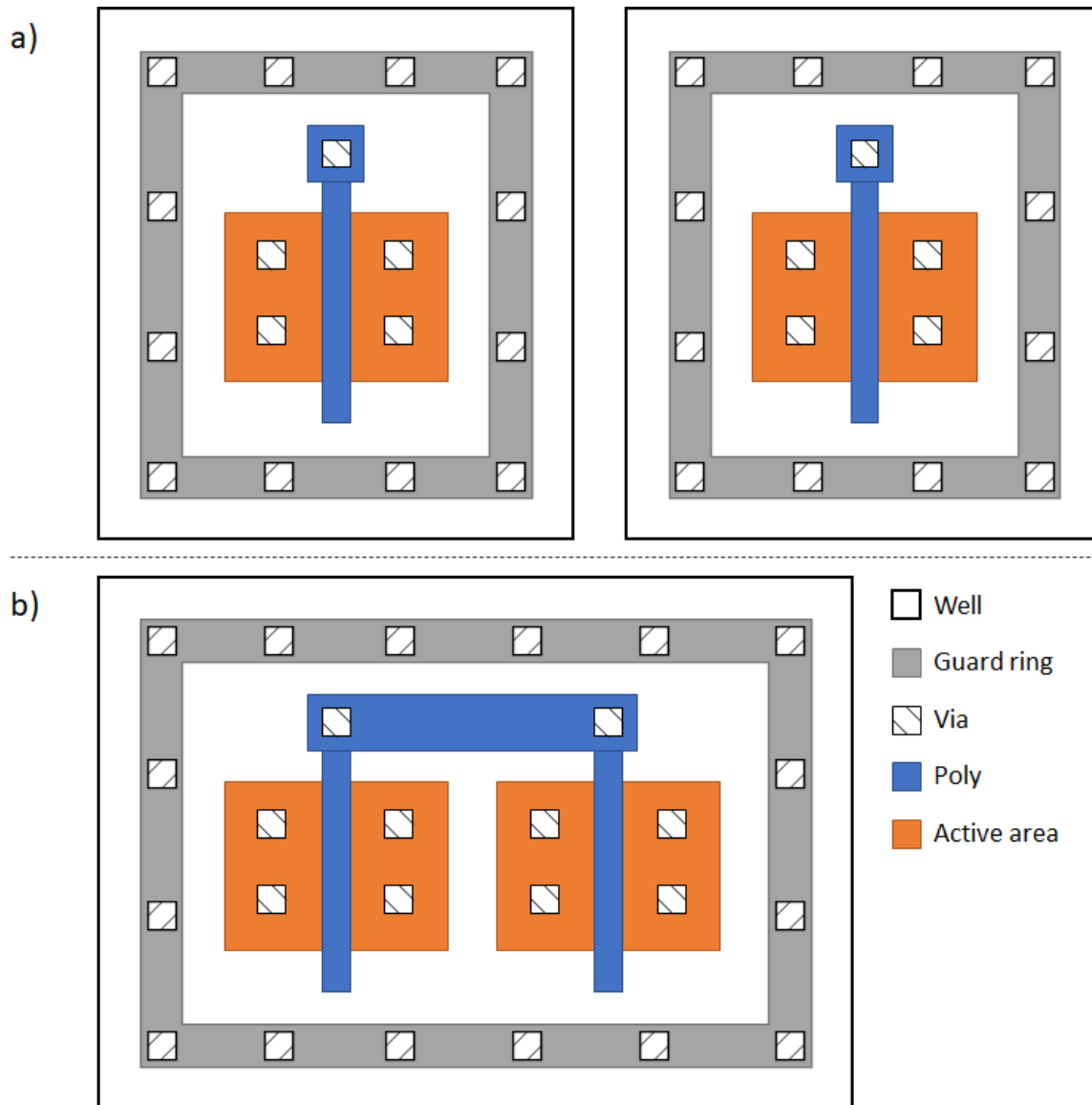


Figure 3.3: Layout diagram displaying transistor isolation. (a) Shows 2 separated simple devices, while (b) shows the same 2 simple devices sharing the same well and guard ring, as well as gate (through poly) and body (through guard ring) connections.

temperature within it. Therefore, a Bipolar Junction Transistor (BJT) has been incorporated to function as an on-chip temperature sensor.

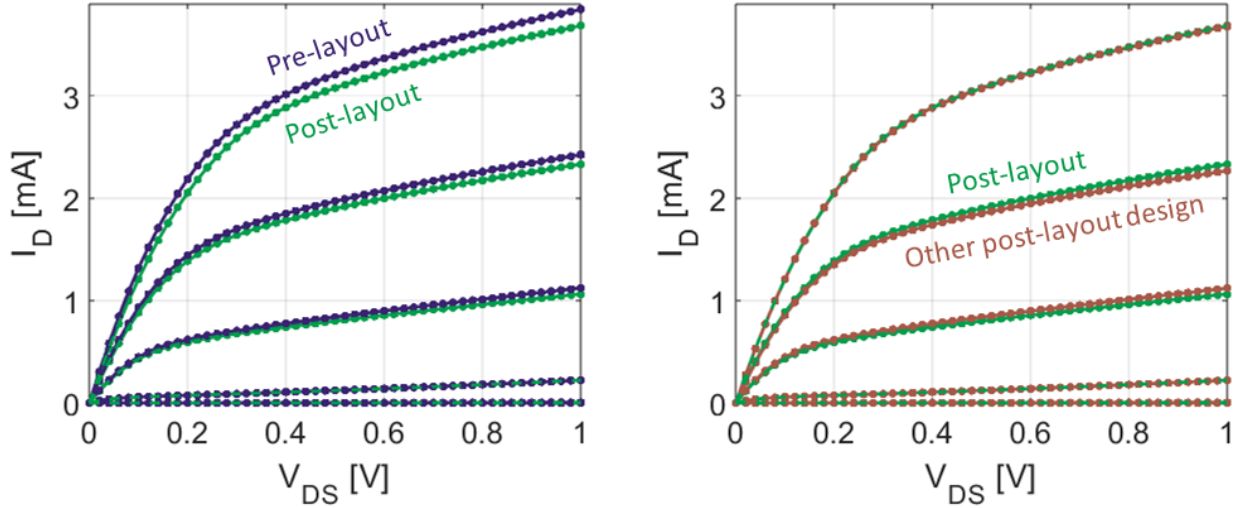


Figure 3.4: Simulated output characteristics of a CMC65n device (a) showing the difference between pre- and post- layout simulations and (b) the difference between two different layout implementations of the same device.

The chip ambient temperature change is measured through the change in the forward bias voltage of the diode-connected BJT (Fig. 3.2b). This relationship is known to be linear at higher temperature range (related to the Fermi-level shift with temperature discussed in section 4.1.3) and not yet tested at cryogenic temperatures. Calibrating the sensor to read absolute cryogenic temperatures is a non-trivial task, considering the unexplored cryo-physics and process variations. Since it is beyond the research scope, the on-chip temperature sensor is not calibrated and used only to determine when the chip temperature has stabilized before initializing any measurement. Moreover, this research is concerned with ambient temperature, and the electron temperature is not measured, due to the nature of this work: modeling the device performance under a certain ambient temperature. Finally, note that these measurements are only possible with the CMC65, and it is assumed that the TSM chips would require similar wait periods.

## 3.2 Cryogenic Measurement Setup

The primary objective of the measurement system is to control the temperature of the DUT while facilitating seamless interfacing between the measurement instruments at RT and the DUT at cryoT. The system should achieve this with minimal additional parasitics (such as resistance) and damage the to DUT. The cryogenic measurements were performed in RAC facilities where

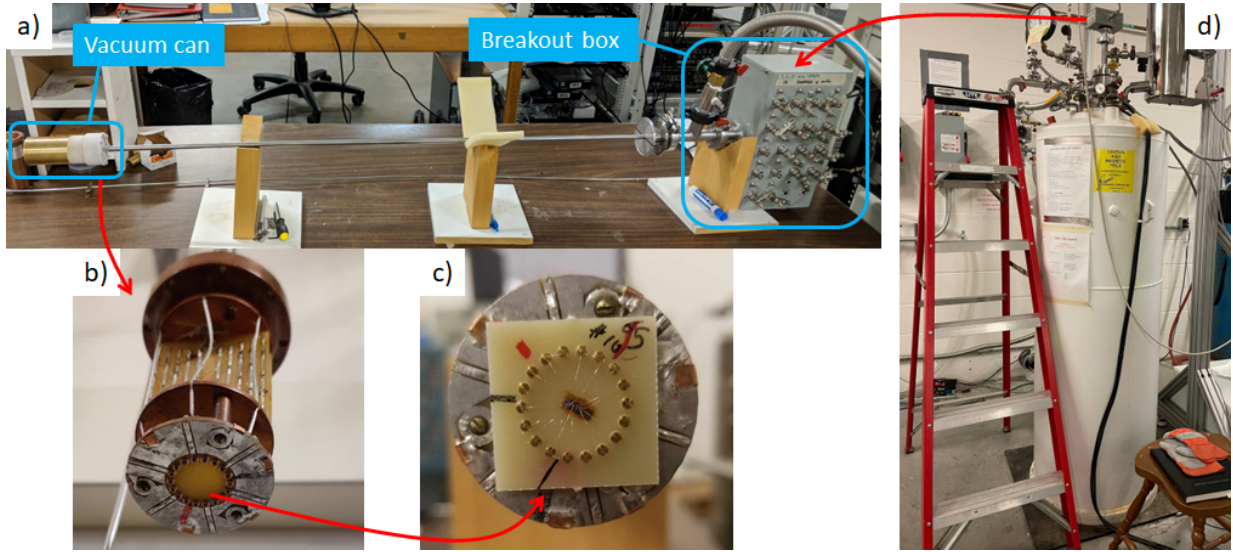


Figure 3.5: Overall view of the cryogenic dipping system showing (a) a sample custom dipper probe, (b) the inside of the mounting vacuum can, (c) a sample wire-bonded to a custom chip holder and mounted into the vacuum can, and (d) the Janis cryostat.

a dipping technique is used (as a cryogenic probing station was not available). In this process, a dipper probe containing the DUT is submerged into the cryogenic medium. A cryogenic probing station was not available. The  $\sim 2m$  dipper probe is comprised of a vacuum can that houses the sample to be tested (at cryoT), a breakout box providing an interface to the DUT (at RT), and a metal rod connecting the two components with enough length to minimise thermal conductivity from the RT side to the cryoT side. After mounting the sample, the air is pumped out of the probe to prevent it from freezing around its internal wiring or the DUT, thus compromising their integrity. Then, the probe is grounded and loaded slowly and carefully into the cryo-dewar to prevent contamination, maintain safe pressure levels, and minimize helium loss.

Initial measurements were performed by submerging the samples in separate helium and nitrogen cryo-dewars, providing only two cryogenic temperature points. Between cryo points, the probe is heated back to RT and then submerged into a new cryo-dewar to remove the condensation formed on it (preventing contamination if helium is next), and to ensure that the sample is not at a cooler temperature than intended (if nitrogen is next). Results published in [28] were measured using this system. However, due to the repeated temperature cycling and extensive handling involved, measurement yield was significantly reduced as samples failed between temperature points (especially TSM samples), resulting in incomplete and inconsistent results.

To combat this problem, a Lake Shore pumped  $4\text{He}$  cryostat with base a temperature of 1.5

K is used. This system, referred to as Janis, offers temperature control range of 1.5 K - 160 K eliminating substantial amount of failure points and greatly improving measurement yield (no failures were recorded between temperature points). After preparing the dipper probe and measuring at RT, the probe is loaded and cooled down to the lowest temperature point ( $\sim 1.5$  K) for safe operation of the cryostat. Then, the probe is heated up to the target temperature point, from 4K to 160K. The measurements presented in section 3.3 are produced using this system.

### 3.2.1 Dipper Probe

The dipper probe utilized in this study is custom made by prof. Baugh's research group. It contains 18 signal lines, of which 6 are coaxial lines exhibiting minimal wire resistance ( $R_{coax} \approx$

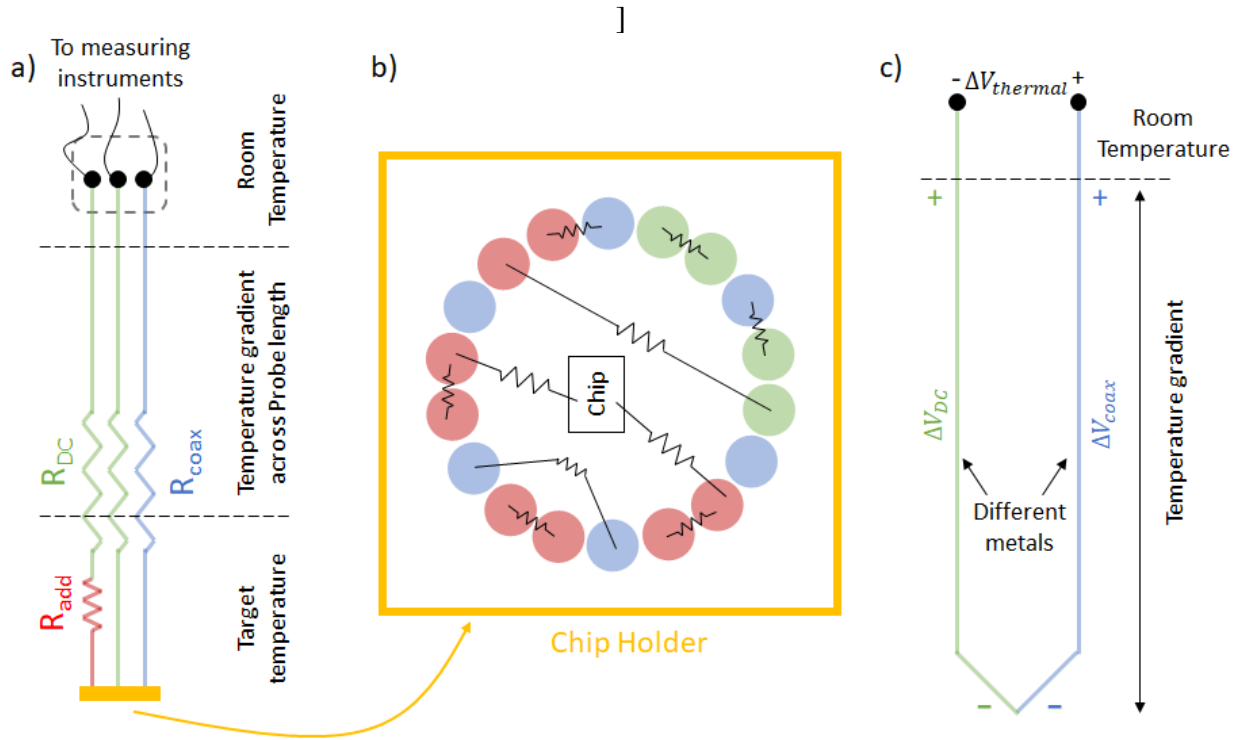


Figure 3.6: Diagram of the internal probe structure where (a) depicts the different line types and resistances, (b) shows the pin arrangement on the chip holder, and (c) illustrates the thermal voltage difference effect. The connections shown in (b) depict a certain test-chip wire-bond map, discussed further in section 3.3, showing parasitic wire-bond resistance.

$2\Omega$  at RT). The remaining lines are twisted pairs DC lines with a wire resistance of  $R_{DC} \approx 120\Omega$  at RT, 8 of them with an additional resistance ( $R_{add} \approx 2.6\text{ k}\Omega$  at RT) resistor, as depicted in Fig. 3.6a. From the sample side, the lines are laid out in a circular pattern as shown in Fig. 3.6b, and the location of each line type is fixed, which must be considered during sample preparation.

Furthermore, it is important to note that a differential voltage can develop when two types of lines (i.e. a coaxial and a DC line) are connected to the same I/O pad. While the exact process is complex, a thermal couple analogy provides enough insight into this phenomenon. A thermal voltage drop  $\Delta V$  will develop across a metal line when put under a temperature gradient. This  $\Delta V$  depends on the line's material properties, and would be different for two types of lines (i.e.  $\Delta V_{coax} \neq \Delta V_{DC}$ ). When the lines are shorted at the cryoT end (having the same potential), the thermal voltage ( $V_{thermal} = \Delta V_{coax} - \Delta V_{DC}$ ) difference will show up at the RT end where measurement instruments are connected.

It is clear that both thermal voltage and probe resistance vary with temperature. In order to mitigate their effect on measurement results, they are measured at each temperature point to be compensated for correctly during measurements (section 3.3).

### 3.2.2 Sample Preparation

The dipping method involves wire-bonding the samples to a custom chip holder compatible with the custom cryogenic dipper probe. During chip preparation, careful planning of each wire-bond (in both electrical and feasibility sense) is necessary while taking into account the probe lines circular layout and fixed placement. For example, the coax lines are mostly desired for source and drain connection (Fig. 3.10), while high resistance DC lines are mostly avoided, but can be used for low DC current signals like gate-voltage.

The TSM samples were intended for precise industry-level testing that favored dense pad placement rather than accessibility, introducing additional wire-bonding challenges. The pads are made small and tightly spaced with side length between  $70\mu m$  and  $35\mu m$ , horizontal pitch between  $125\mu m$  and  $70\mu m$ , and vertical pitch between  $\sim 70\mu m$  to  $\sim 35\mu m$  (almost no vertical spacing) for TSM65 and TSM20, respectively (Fig. 3.7a). To put these dimensions in perspective, wire-bonders typically use  $\sim 25\mu m$  threads that expand even more when mashed down to bond with pads (Fig. 3.7b). The TSM pads are  $>3\times$  than a typical pad size of  $100\mu m \times 100\mu m$ , a dimension significantly larger than wire-bonding threads.

To improve wire-bonding consistency, a number of wire-bonder types, techniques, and thread material were tested. Reliable results were achieved using a wedge wire-bonder with aluminum thread, in contrast to a ball bonder and gold thread. Additionally, a non-standard bonding technique is developed where thread termination happens at the chip side instead of the holder

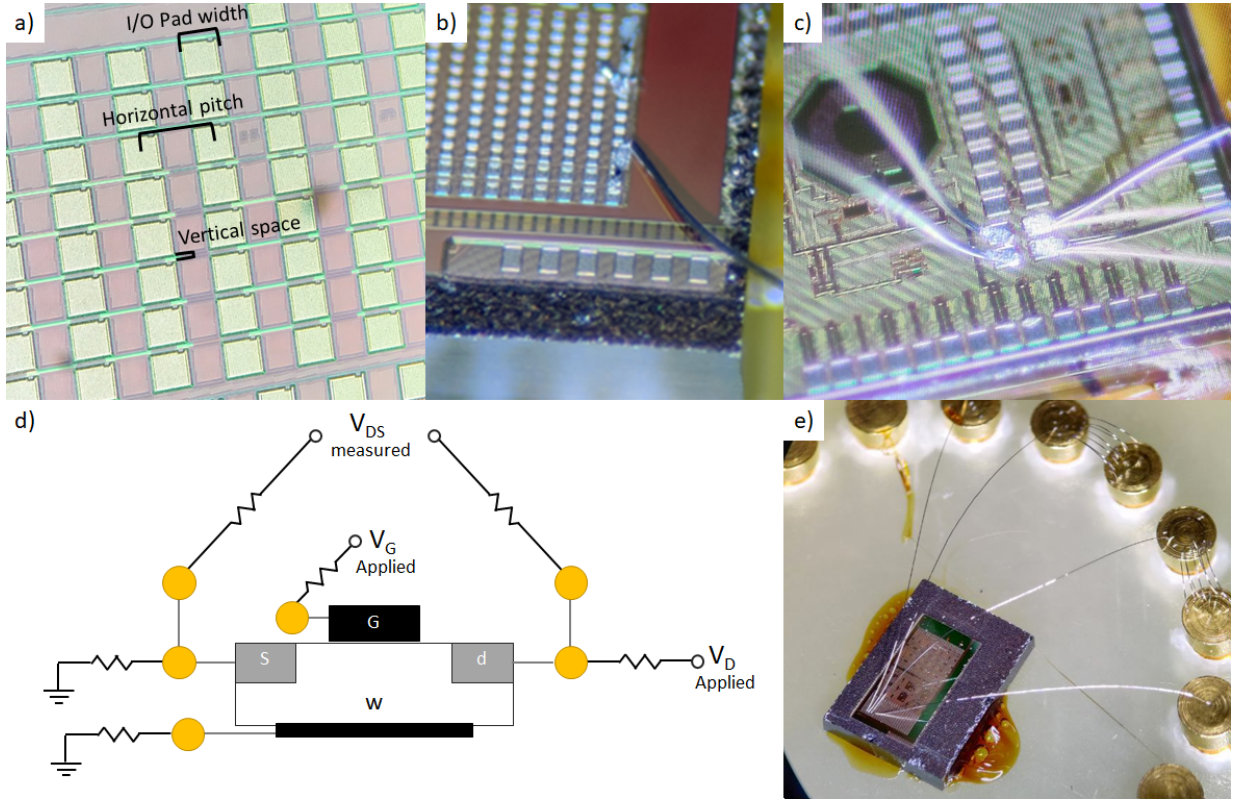


Figure 3.7: (a) Example of a TSM chip with demonstrating relevant feature sizes. (b) Early wire bond attempt, showing the size discrepancy between the thread and pad dimensions. (c) An example of a successful wire bond of a CMC chip showing controlled tail length and wedge size. (d) Typical DUT wiring diagram to differentiate between applied and measured terminal voltages. (e) A successful wire-bonding example of a TSM sample.

side. This is done to reduce the residual tail length (Fig. 3.7b and c) and prevent unintentional connection and short-circuits with neighboring pads or other wire-bonds.

Finally, probe resistance can significantly affect measurement results as shown in Fig. 3.8. It is easy to assume that the entire supplied voltage is seen by the device. and then draw conclusions regarding performance parameters from Fig. 3.8a. However, this perspective hides the additional probe characteristics that must be removed or compensated for during DUT analysis. Fig. 3.8b shows the difference between the two perspectives, and how the device is not driven to its full potential when turned-on. From  $V_{DS}$  perspective, the resistance of both source and drain lines are added together, while  $V_{GS}$  is only affected by the source's line, since the gate current is small, and

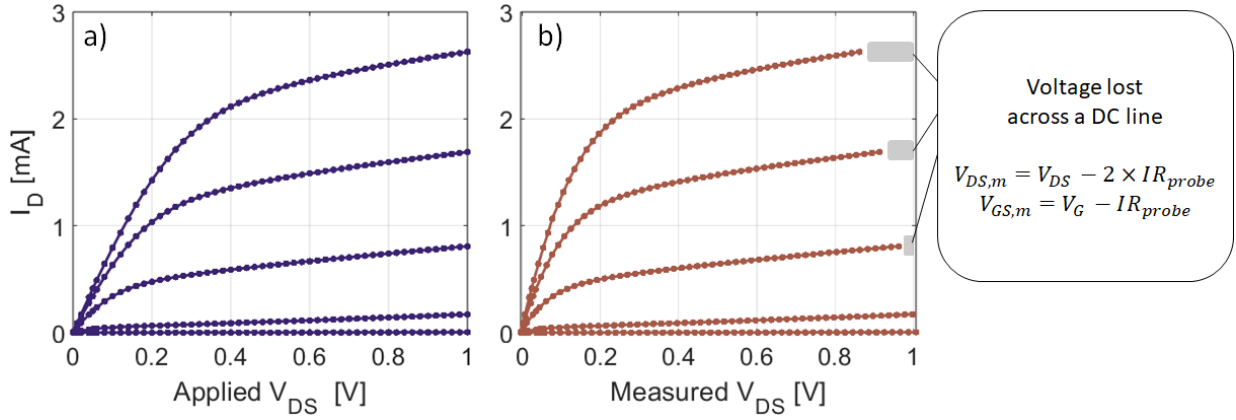


Figure 3.8: Output characteristics of a TSM65n sample where the current is plotted against (a) the applied voltage, and (b) the measured voltage across the source and drain terminals (as shown in Fig. 3.7d). The gap shows the significance of the of voltage lost across the probe resistance manifested purely by the line's length (no resistors added).

the voltage lost across the gate line is negligible. For simplicity, the same type of line is always used for the source and drain. Therefore, it is desired to use the low resistance coax lines ( $\sim 4\Omega$ ) instead of the orders of magnitude resistance DC lines ( $\sim 240\Omega$  or  $4k\Omega$ ). The measurements shown from this point onward will be from the applied  $V_{DS}$  voltage perspective, and this effect will be compensated for during analysis.

### 3.2.3 Experimental Setup

Regarding the electrical connections, a single MOSFET device requires a connection to each of its 4 terminals (more if ESD cells are involved). It was not always possible to use the six coax lines exclusively (low resistance lines), especially when multiple devices are wired in the same chip simultaneously. Moreover, the circular pattern of the chip-holder pins and fixed placement of line types (according to Fig 3.6b) hinders the flexibility of choosing line types. Hence, it is necessary to use a 4-point measurement (Fig.3.7d) at the source and drain terminals if the DC lines are used. This is to discern the difference between the applied voltage at the breakout box and the measured voltage across the DUT terminals. This technique is not used on the body and gate terminals since the small expected DC current would produce negligible voltage drop across the line. 4-point measurements increase the minimum number of lines required per device and further complicates wire-bonding. Only a maximum of two isolated devices (usually a PMOS and an NMOS) were successfully wired and measured simultaneously.

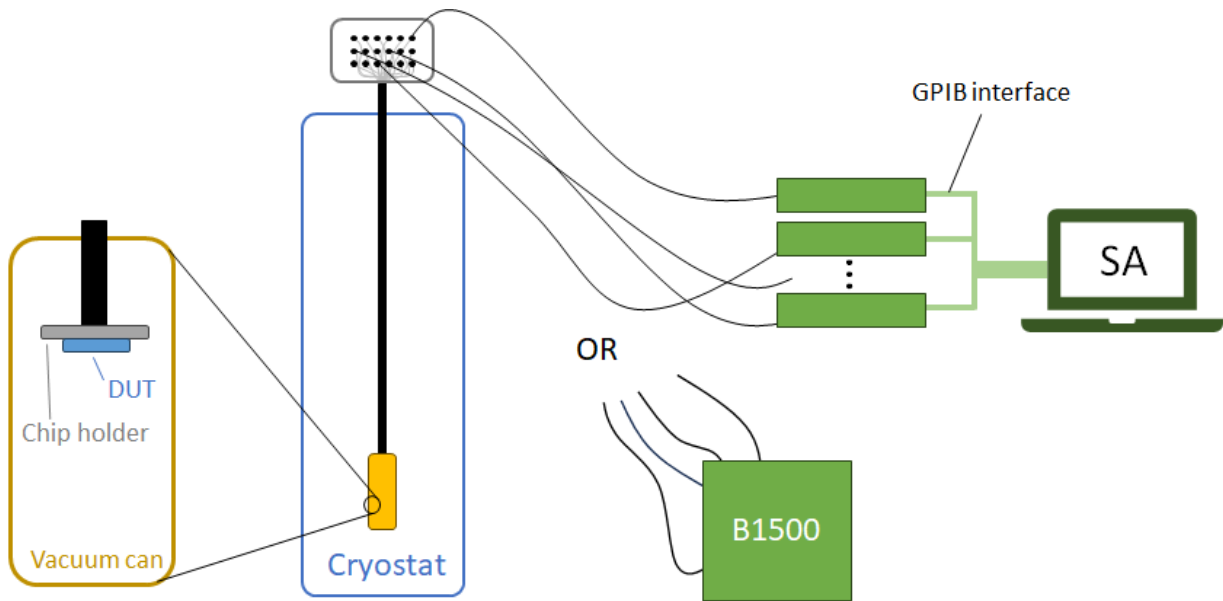


Figure 3.9: Diagram of the overall measurement setup, showing two possible setups: the self-contained B1500 semiconductor device analyzer, or a group of individual instruments (digital multi-meters, SMUs, power sources, etc) connected to an external machine with SA.

For the on-ship temperature sensor present on the CMC65 samples, an external  $10k\Omega$  –  $100k\Omega$  resistor is used (Fig. 3.2b). The exact resistor value is not critical given that only temperature change is measured, but this range is chosen to control the current levels. The connection of the BJT involves the base and collector terminals being tied together, while the emitter and substrate terminals are connected to ground, requiring only two lines. With transistor values not critical, even the high-resistance DC lines can be used. This frees up the coax and low resistance DC lines for more critical measurements.

A series of measurement instruments are connected to the DUT through the exposed break-out box and controlled by an in-house (RAC) developed data acquisition software called Spanish Acquisition (or SA). Using the IEEE-488 GPIB interface allowed for great level of flexibility and modality. The system was updated to support remote control around mid 2022, which substantially improved the ability conduct measurements efficiently. As shown later in Fig. 3.10, a significant amount of time is required for the chip to settle at most of the temperature points. Combined with the nighttime curfew in the lab, a single temperature cycle would require 4-5 days, which introduced further equipment booking and scheduling challenges. Remote control allowed for overnight measurements and blocked the cryostat for less than 2 days.

Unfortunately, this system is limited by the 2 available SMUs. Therefore, a Keysight B1500 semiconductor analyzer with 3 SMUs is occasionally utilized when gate and substrate currents are measured. Simple multi-meters are not suitable to measure such low currents. The B1500 is not used permanently because of its proprietary software that limits modularity, stimulus setup flexibility, remote access. For example, the 3 included SMUs are not enough for 4-point measurements, requiring the exclusive use of the very low resistance coax lines.

### 3.3 Experimental Results

It is important to distinguish the probe's influence on the measurements and compensate for it when analysing the DUT. Therefore, the following subsections will divide the experimental results into two categories: probe characteristics and DUT characteristics.

#### 3.3.1 Dipping Probe Characteristics

It is imperative to characterize the probe parasitics and study their temperature variations to be able to separate them from the DUT's own temperature variation (the intended target of the

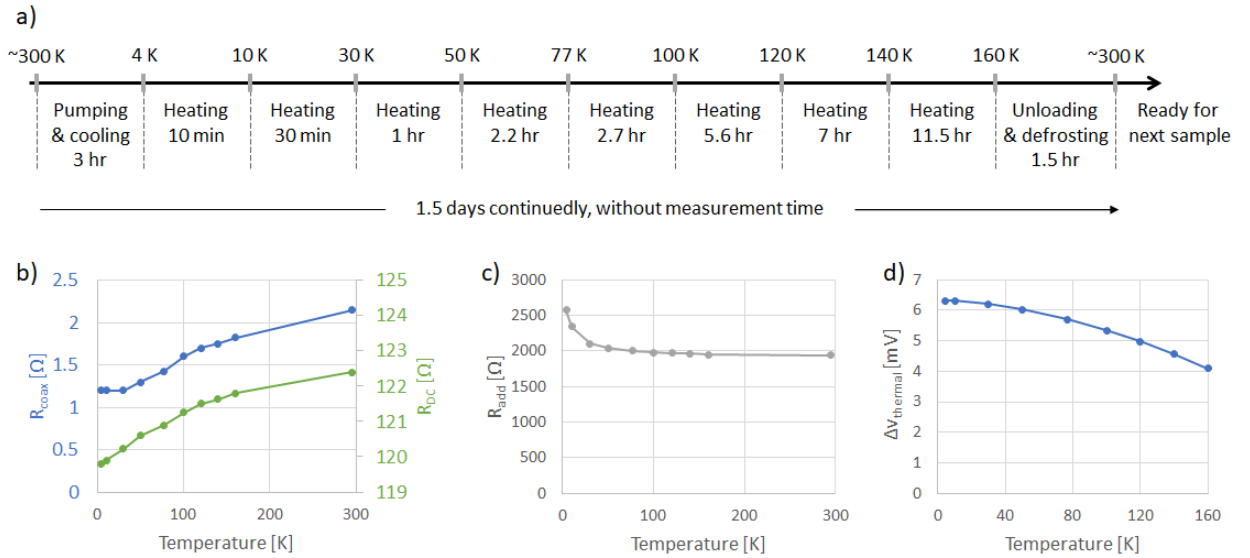


Figure 3.10: Dipping probe characteristics. (a) Wait time period required between temperature points. (b-c) Line resistance and (d)  $\Delta V_{thermal}$  variation with temperature.

experiment). To this end, a series of short circuits, along with the temperature sensor circuit (Fig. 3.6b), have been employed to determine the various line resistances and  $\Delta V_{thermal}$  at each cryogenic temperature point. Since the cold side cannot be probed directly, two lines of the same type are connected together, and the line resistance is divided by two. The wire-bond thread resistance  $R_{thread}$  is measured directly at RT to be less than  $0.1\Omega$ . This is repeated for multiple wire-bond lengths, indicating that the correct value is too small for a handheld digital multi-meter to measure. Given the order of magnitude of  $R_{thread}$ , it is not measured further to a precise value and is ultimately ignored. Additionally, time required for the sample to transition between temperature points has also been recorded.

The results of the characterization, presented in Fig. 3.10, indicate a substantial increase in  $R_{add}$  by 33% at 4K, whereas line resistances of  $R_{DC}$  and  $R_{coax}$  decreased with temperature by  $2.6\Omega$  and  $< 1\Omega$ , respectively, a negligible change. The thermal voltage is measured between every line pair combination and, as expected, it only develops when coax and DC lines are shorted together at the cold end. It showed consistent increase with decreasing temperature, at a very small voltage ( $< 10mV$ ). Nevertheless, the probe resistance and offset voltage is compensated for during post data processing.

### 3.3.2 Device Characteristics

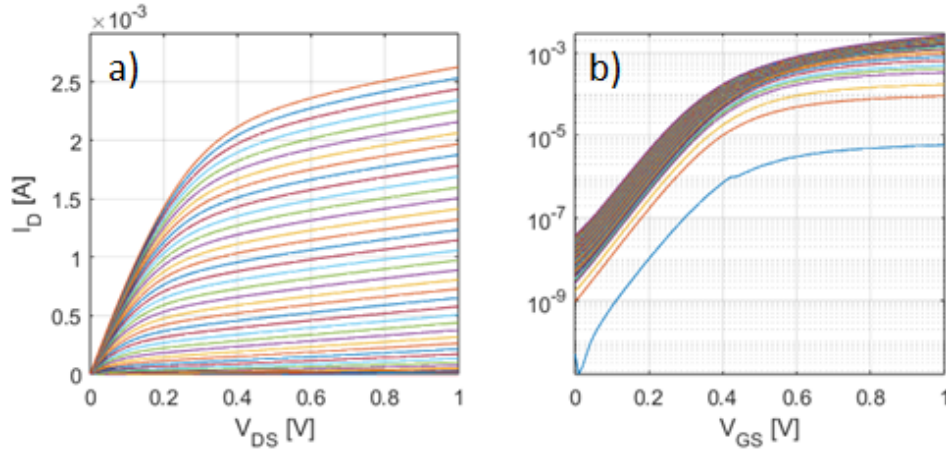


Figure 3.11: Pre-processing experimental results from a CMC65n device at RT. Both plots present the same data from two different perspectives: (a) output and (b) transfer characteristics. The voltage biases used for both  $V_{GS}$  and  $V_{DS}$  are:  $0 \rightarrow 1$  with a step size of 20 mV, in addition to 10 mV and 50 mV.

To extract the DC characteristics of MOSFET devices, two measurements are needed: the output characteristics and the transfer characteristics. The former involves sweeping  $V_{DS}$  over many points for a few  $V_{GS}$  values, while the latter requires the opposite. If conducted back-to-back, the overlapping ( $V_{GS}$ ,  $V_{DS}$ ) points are likely to have different measured currents due to noise fluctuations or lingering effects from previous measurements (like self-heating, or aging). Moreover, signs of reduced reliability have been reported in literature [65] and observed during testing. Therefore, a "one-measurement per temperature-point" approach is adopted, where both  $V_{GS}$  and  $V_{DS}$  are swept to form a high resolution 2-dimensional matrix. Additionally, to mitigate any potential temperature cycle effects on device performance, only the first device measured per chip is included. The resulting dense data is visualized as many  $I_D$ - $V_{DS}$  and  $I_D$ - $V_{GS}$  curves as seen in Figs. 3.11. The data is then post-processed to the perspective desired via an in-house developed software (chapter 5) producing the characteristics shown in Figs. 3.12 and 3.13. With regards to CMC65, it has been determined that ESD protection was not needed, and hence only unprotected devices are shown to eliminate any unnecessary interference from ESD protection.

The overall increase in drive current with decreasing temperatures is apparent in the output characteristics of most devices. The saturation current, however, seems to decrease for low  $V_{GS}$  values, attributed to the notable increase in the threshold voltage. For example, Fig. 3.12a displays current increase for  $V_{GS} = 1$  V curves, but almost no change visible for  $V_{GS} = 0.8$  V, and decreasing current for lower  $V_{GS}$  curves. This effect is most pronounced in the TSM20n device (Fig. 3.12g) that has the most apparent threshold voltage at RT (close to the maximum operating voltage, Fig. 3.12h) and does not display any current increase at cryoT. The transfer characteristics for all devices show monotonic improvement in subthreshold swing  $S$  (visible as steeper slope) and steady increase in threshold voltage (noticeable as a shift to the right). The  $V_{GS}$ -dependence of subthreshold swing is clearly visible in all devices at cryoT, especially in the TSM20 devices Figs. 3.12h and 3.13h. Finally,  $S$  is extracted from the linear transfer characteristics ( $V_{DS} = 0.05$  V) using the constant current method (where the slope is measured at the same  $I_D$  level across all transfer curves, even if that falls on different  $V_{GS}$  biases). The current levels are chosen carefully to avoid the effect of Coulomb blockade on the subthreshold region and simplify  $S$  extraction.  $S$  saturation behaviour is especially pronounced in TSM28p and TSM20p with noticeably low saturation temperature.

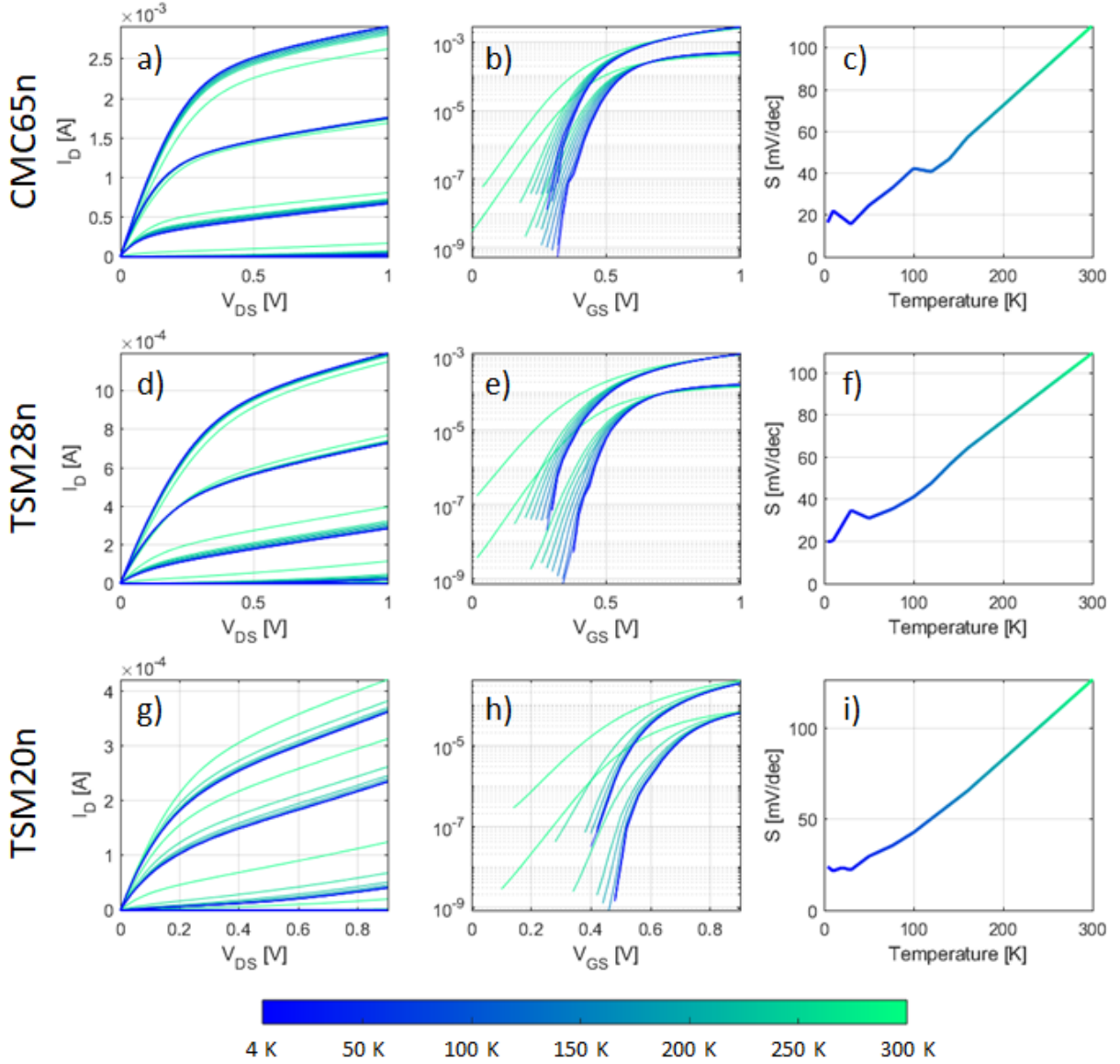


Figure 3.12: Measured characteristics of NMOS bulk devices manufactured in (a, b, c) 65-nm, (d, e, f) 28-nm, and (g, h, i) 20-nm processes, with channel length of 65 nm, 30 nm, and 20 nm, respectively. In (a, d, g)  $V_{GS} \in [0, 1]$  with 0.2V step, except for TSM20 where the maximum operating voltage is 0.9V. In (b, e, h)  $V_{DS} = [0.05, 0.8]$ V, and  $S$  in (c, f, i) is measured at  $V_{DS} = 0.05$ V. Note that some  $I_D$ - $V_{DS}$  curves corresponding to low  $V_{GS}$  biases may overlap each other, giving the illusion that only four  $V_{GS}$  points are displayed.

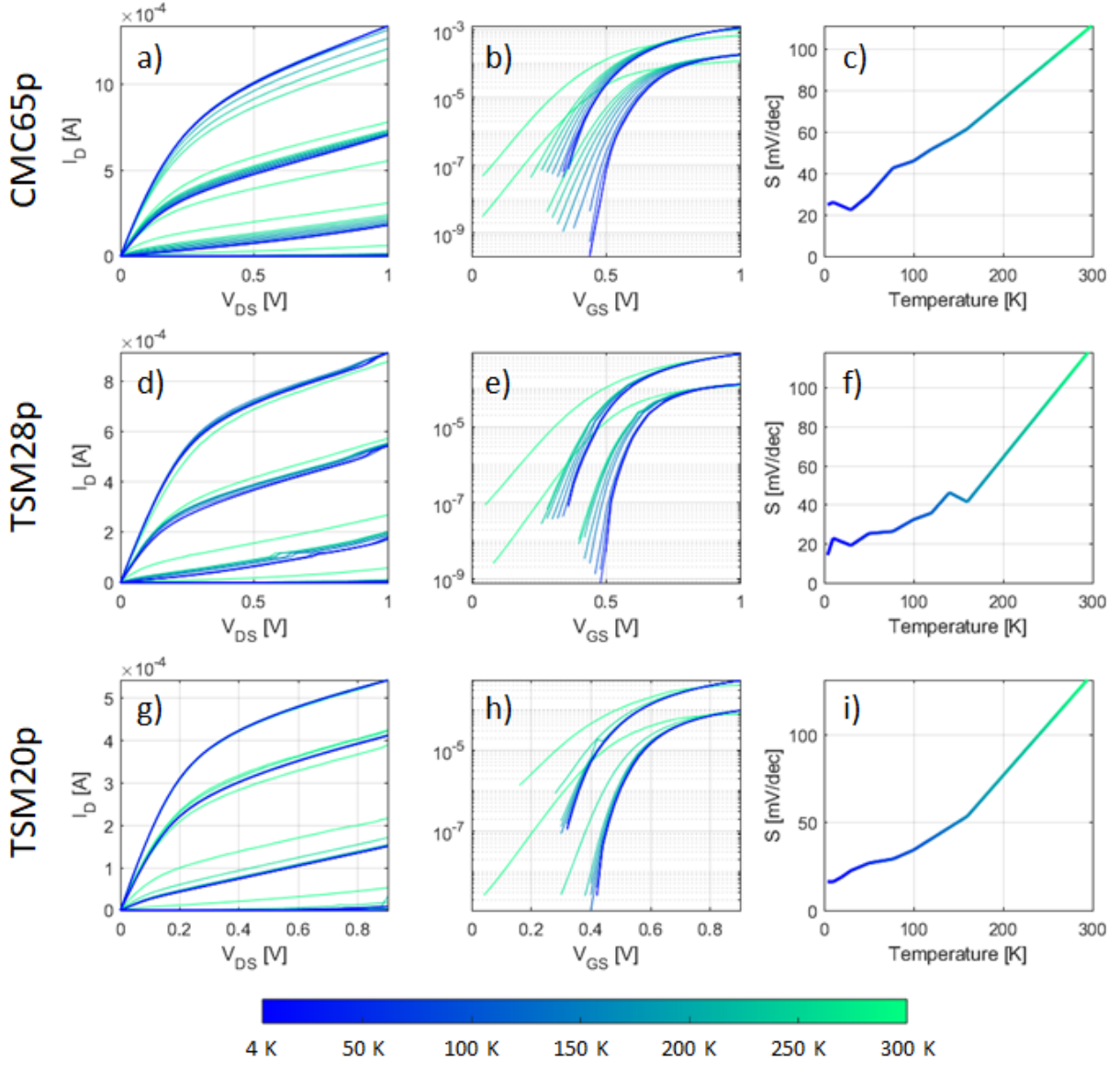


Figure 3.13: Measured characteristics of PMOS bulk devices manufactured in (a, b, c) 65-nm, (d, e, f) 28-nm, and (g, h, i) 20-nm processes, with channel length of 65 nm, 30 nm, and 20 nm, respectively. In (a, d, g)  $V_{GS} \in [0, 1]$  with 0.2V step, except for TSM20 where the maximum operating voltage is 0.9V. In (b, e, h)  $V_{DS} = [0.05, 0.8]V$ , and  $S$  in (c, f, i) is measured at  $V_{DS} = 0.05V$ . Note that some  $I_D$ - $V_{DS}$  curves corresponding to low  $V_{GS}$  biases may overlap each other, giving the illusion that only four  $V_{GS}$  points are displayed.

# Chapter 4

## Cryogenic MOSFET

In the preceding chapters, it has been established that operating MOSFETs at extremely low temperatures leads to unforeseen changes in their behavior. These alterations surpass the capabilities of current-day commercial compact models due to the modified physical properties of cryogenic semiconductors and resulting changes in operation mechanisms within the transistor device. Consequently, the development of a cryoMOSFET compact model necessitates a reassessment of the underlying principles of semiconductor physics at cryoT. It is imperative to investigate the suspected governing phenomena, including incomplete ionization, bandgap widening, interface traps, band-tail states, and others. By analyzing the device physics in light of these new findings, it becomes possible to simplify and utilize these effects for the development of a flexible cryogenic MOSFET compact model to effectively represent a wide range of device technologies using only a few fitting parameters. The availability of such a model is essential for the design and optimization of MOSFET-based electronic circuits intended for low-temperature applications.

### 4.1 Device Physics

The physical properties of semiconductors are often studied at room temperature, benefiting from the ability to make certain assumptions and employ useful approximations. However, when confronted with extreme temperatures, this analysis must be undertaken without relying on those simplifications, and implications that stem from their possible invalidity must be considered. By unveiling the governing phenomena that dictate semiconductor behavior at extremely low temperatures, a comprehensive understanding of cryogenic MOSFET operation can be achieved.

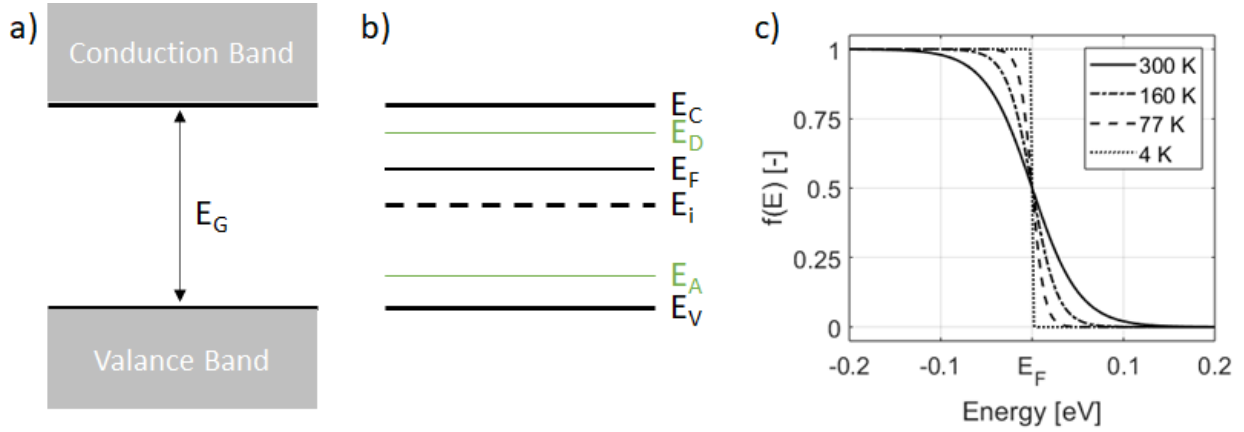


Figure 4.1: Visualization of the energy band model (energy vs. spacial dimension) showing (a) the three bands and (b) main energy levels. The semiconductor shown in (b) is an n-type since  $E_F$  is within the upper half of the band. (c) The Fermi-Dirac distribution function under multiple temperatures, displaying the transition from a smooth function over a wide range of energies ( $\sim 0.3$  eV) at RT to a sharp step function at deep cryoT.

The following discussion will focus heavily on a fundamental concept in solid-state physics and semiconductor device engineering known as the energy band model. The widely used model provides a visual representation of the energy levels and band structure within the semiconductor, delivering valuable insights into charge carrier behavior, charge transport mechanisms, and the response to external stimuli like electric fields and light. The energy band model is applicable at cryoT.

Fig. 4.1a illustrates the basics of the model showing five key energy levels within three main regions known as bands. The valence band represents the energy levels where chemically bound electrons reside, while the conduction band represents the unoccupied energy levels where an electron would be free to move and conduct current. For an electron to break free from its bond and contribute to current flow, it has to gain energy equal to the band gap ( $E_G$ ) to move from the top of the valance band ( $E_V$ ) to the bottom of the conduction band ( $E_C$ ). Regardless of the existence of states at a given energy, the probability of carrier occupation is governed by the Fermi-Dirac (FD) distribution function:

$$f(E) = \frac{1}{1 + e^{\frac{E-E_F}{k_B T}}} \quad (4.1)$$

where  $E_F$  is the Fermi energy level, arguably the most important energy level in Fig.4.1b, marking the 50% probability of occupation at that energy level. As illustrated in Fig. 4.1c, the

temperature ( $T$ ) exercises significant influence over the function's smoothness and range. The exponential influence of  $T$  is one of the reasons why semiconductor physics must be carefully re-examined at cryoT. In an undoped semiconductor (AKA intrinsic), this energy level is given a special name: the intrinsic Fermi level ( $E_i$ ) where the electron density in the conduction band is equal to the hole (or electron vacancy) density in the valance band.

The last two energy levels shown in Fig.4.1b are the donor ( $E_D$ ) and acceptor ( $E_A$ ) levels, introduced to the semiconductor when it is intentionally injected (or doped) with impurities to manipulates its electrical properties. The additional electrons contributed by donors (or holes created by acceptors) reside on  $E_D$  or  $E_A$ , exhibiting ionization energies significantly smaller than  $E_G$ . Consequently, the carrier occupation probability is affected, and  $E_F$  shifts closer to  $E_C$  (in an n-type semiconductor) or  $E_V$  (in a p-type semiconductor), depending on the excess electron or hole concentrations, respectively.

In the following subsections,  $E_F$  will play a central role for most analysis, where its location within the band gap provides ample information regarding the semiconductor properties.

#### 4.1.1 Field-Assisted Ionization

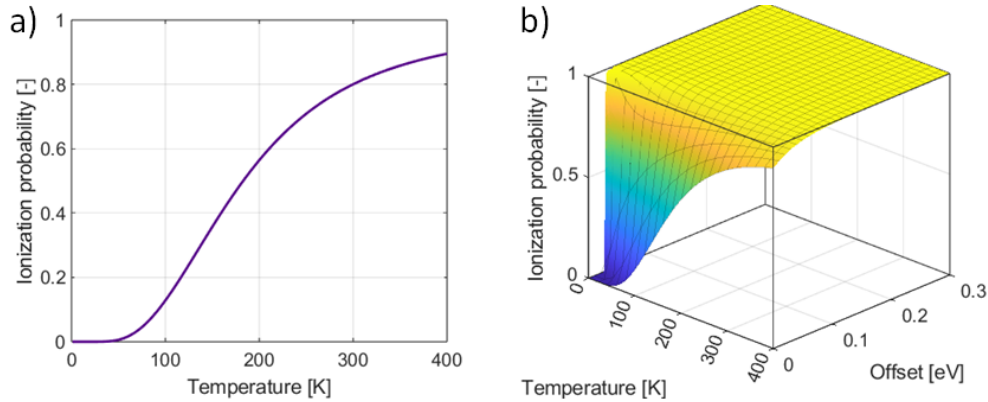


Figure 4.2: Probability of dopant ionization in a semiconductor vs.  $T$ . (a) shows the 2-dimensional relationship while (b) adds energy offset as a third dimension. These results are generated using e.q. 4.1, where  $E_A$  set 45 meV above  $E_V$ .

Before diving any deeper, one might rightfully wonder: what about freeze-out (the term for when  $T$  is too low for dopants to undergo ionization, i.e. become the free carriers that facilitate current flow)? How do FET devices operate at these conditions when there are no free carriers

to carry current (Fig. 4.2a)? To answer these questions, one must analyze the effects of applying various voltage biases across the device's terminals.

Fig. 4.2a shows the ionization probability in the absence of any applied voltages (Fig. 4.1b) calculated using e.q. 4.1 at  $E_A$  or ( $f(E_A)$ ). Indeed, negligible to nonexistent free carrier concentration is expected at temperatures below  $\sim 50$  K. However, during their operation, devices typically have various voltages applied across their terminals, creating electric fields within the device, and shifting/bending the bands around  $E_F$ . This can be modeled by a simple offset added to e.q. 4.1:

$$f(E_A) = \frac{1}{1 + e^{(E_A - E_F + offset)/k_B T}} \quad (4.2)$$

This adds a third dimension to the ionization probability plot (Fig. 4.2b) revealing the field-assisted ionization phenomena where the ionization probability quickly approaches unity even for a small amounts of offset. This shows that under these conditions it is not unreasonable to assume significant ionization, or even complete ionization of dopants. Thus, device operation is possible even under stringent device physics analysis at cryoT.

However, there is a lot more to understand about this operation, and many other convenient assumptions must be re-visited.

## 4.1.2 Fermi Level Location

The validity of the two most common simplifications made during RT semiconductor analysis (complete ionization and semiconductor non-degeneracy) must be investigated down to deep cryoT. According to the charge-neutrality principle, the total charge density ( $\rho$ ) in a semiconductor under thermal equilibrium is zero. Therefore:

$$\rho = N_D^+ - N_A^- + p - n = 0 \quad (4.3)$$

where  $N_D^+$  and  $N_A^-$  are the ionized concentrations of donor and acceptor impurities, respectively, while  $n$  and  $p$  are the electron and hole concentrations in the conduction and valance bands, respectively. It is critical to differentiate between the total and ionized dopant concentrations because while the total concentration provides a measure of the overall doping level in the material, it is the ionized dopants that manipulate carrier concentrations within the material and shift  $E_F$  to the top or bottom of the band gap. The first common simplification comes from assuming complete ionization, where all injected impurities gain or lose electrons and the full potential of semiconductor manipulation is realized. A more conservative approach is to calculate the portion of ionized impurities, regardless of it being 100% or lower, as shown in Table 4.1. The equations

Table 4.1: Ionized impurities equations

|           | Complete ionization                                | Incomplete ionization                   |
|-----------|----------------------------------------------------|-----------------------------------------|
| $N_D^+ =$ | $N_D$ , or $\frac{N_D}{1+g_D e^{(E_F-E_D)/k_B T}}$ | $\frac{N_D}{1+g_D e^{(E_F-E_D)/k_B T}}$ |
| $N_A^- =$ | $N_A$ , or $\frac{N_A}{1+g_A e^{(E_A-E_F)/k_B T}}$ | $\frac{N_A}{1+g_A e^{(E_A-E_F)/k_B T}}$ |

in the table use the degeneracy factors  $g_A$  and  $g_D$  to refer to how many carriers a dopant implant site can accommodate, unrelated to semiconductor degeneracy nor  $E_F$  location.

$N_D^+$  and  $N_A^-$  influence  $E_F$  location,  $p$  and  $n$  are determined by  $E_F$  and the effective density of valance and conduction band states ( $N_C$  and  $N_V$ , respectively). However, obtaining an exact solution for  $n$  and  $p$  involves the utilization of the complex FD integral,  $\mathcal{F}_n(\eta)$ , which is often circumvented by employing the Maxwell-Boltzmann (MB) approximation, the second of the common simplifications made during RT analysis (Table 4.2). This simplification comes with one condition:  $E_F$  does not lie within  $3k_B T$  from the band gap edges, or the semiconductor is degenerate and the approximation is inaccurate. Therefore, it is only valid for non-degenerate semiconductors, which are limited to doping concentrations of  $\sim 1 \times 10^{18} \text{ cm}^{-3}$  in Si at RT. Nonetheless, it is important to acknowledge that this limit may vary under different temperature conditions, warranting further investigation.

Table 4.2: Carrier concentration equations

|       | Degenerate semiconductor                                       | Non-degenerate semiconductor                                                                                         |
|-------|----------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|
| $n =$ | $N_C \mathcal{F}_{1/2} \left( \frac{E_F - E_C}{k_B T} \right)$ | $N_C e^{\left( \frac{E_F - E_C}{k_B T} \right)}$ , or $N_C \mathcal{F}_{1/2} \left( \frac{E_F - E_C}{k_B T} \right)$ |
| $p =$ | $N_V \mathcal{F}_{1/2} \left( \frac{E_V - E_F}{k_B T} \right)$ | $N_V e^{\left( \frac{E_V - E_F}{k_B T} \right)}$ , or $N_V \mathcal{F}_{1/2} \left( \frac{E_V - E_F}{k_B T} \right)$ |

$\mathcal{F}_{1/2}(\eta) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{x^{1/2} dx}{1+e^{x-\eta}}.$

Finding an  $E_F$  value that satisfies charge-neutrality (eq. 4.3) at cryoT requires specific knowledge of the ionization degree and semiconductor degeneracy before employing any of the equation from Tables 4.1-4.2 to find the needed concentrations. To this end, the strategy is to use the most inclusive forms (FD statistics and incomplete ionization) to avoid making any implicit assumptions. The results are then utilized to discern which simplifications can be made at which temperatures.

For an n-type semiconductor, acceptors can be assumed absent ( $N_A = N_A^- = 0$ ), leading to the

following relationship:

$$N_V \mathcal{F}_{1/2} \left( \frac{E_V - E_F}{k_B T} \right) - N_C \mathcal{F}_{1/2} \left( \frac{E_F - E_C}{k_B T} \right) + \frac{N_D}{1 + g_D e^{(E_F - E_D)/k_B T}} = 0 \quad (4.4)$$

$$N_V \mathcal{F}_{1/2} \left( \frac{E_V - E_F}{k_B T} \right) - N_C \mathcal{F}_{1/2} \left( \frac{E_F - E_C}{k_B T} \right) - \frac{N_A}{1 + g_A e^{(E_A - E_F)/k_B T}} = 0 \quad (4.5)$$

It is worth noting that a dependency loop exists between the donor and carrier concentrations. The value of  $N_D^+$  must match  $n$ - $p$ , while the values of  $n$  and  $p$  are dictated by  $E_F$ , which is in turn determined by  $N_D^+$ . Deriving an analytical solution for this relationship to find  $E_F(T)$  is a complex task due to the presence of the FD integral. Instead, a numerical solution is calculated using MATLAB for multiple doping concentrations across a wide range of temperatures, all the way to 1 K.

For simplicity, the analysis assumes Boron (B) for acceptor and Phosphorus (P) for donor impurities because they share the same ionization energy:  $E_C - E_D = E_A - E_V = 45$  meV [71]. Moreover, the analysis also accounts for the band gap widening phenomenon, where the band gap width increases with decreasing temperature, is modelled as [70]:

$$E_G(T) = E_G(0) - \frac{\alpha T^2}{T + \beta} \quad (4.6)$$

## Semiconductor Degeneracy Level

Fig. 4.3 depicts two competing phenomena for degeneracy: the shifting  $E_F$  towards the band edges risking degeneracy, and the shrinking degeneracy threshold ( $E_C - 3k_B T$  or  $E_V + 3k_B T$ ) as temperature decreases, making it harder to reach degeneracy. For the conditions simulated (Si uniformly doped with B or P) Fig. 4.3a shows that non-degenerate semiconductors at RT remain non-degenerate down to deep cryoT, allowing for MB approximation to take place across all temperatures. A deeper look reveals that a small range of doping concentrations (between  $2 \times 10^{18} \text{ cm}^{-3}$  and  $2.5 \times 10^{18} \text{ cm}^{-3}$ , Fig. 4.3c) briefly crosses the degeneracy threshold before crossing back into non-degeneracy. Nevertheless,  $E_F$  does not stray from the degeneracy threshold by a significant amount, and the MB approximation remains accurate.

Moreover, degenerate doping levels above  $2.5 \times 10^{18} \text{ cm}^{-3}$  becomes non-degenerate at deep cryoT. Fig. 4.3a shows that as  $T \rightarrow 0$   $E_F$  clearly converges to the midpoint between the band edge and the corresponding dopant energy level ( $E_F \rightarrow (E_C + E_D)/2$  or  $(E_A + E_V)/2$ ). This observation corroborates the MB-based analysis illustrated in [71], further validating the approximation at deep cryoT.

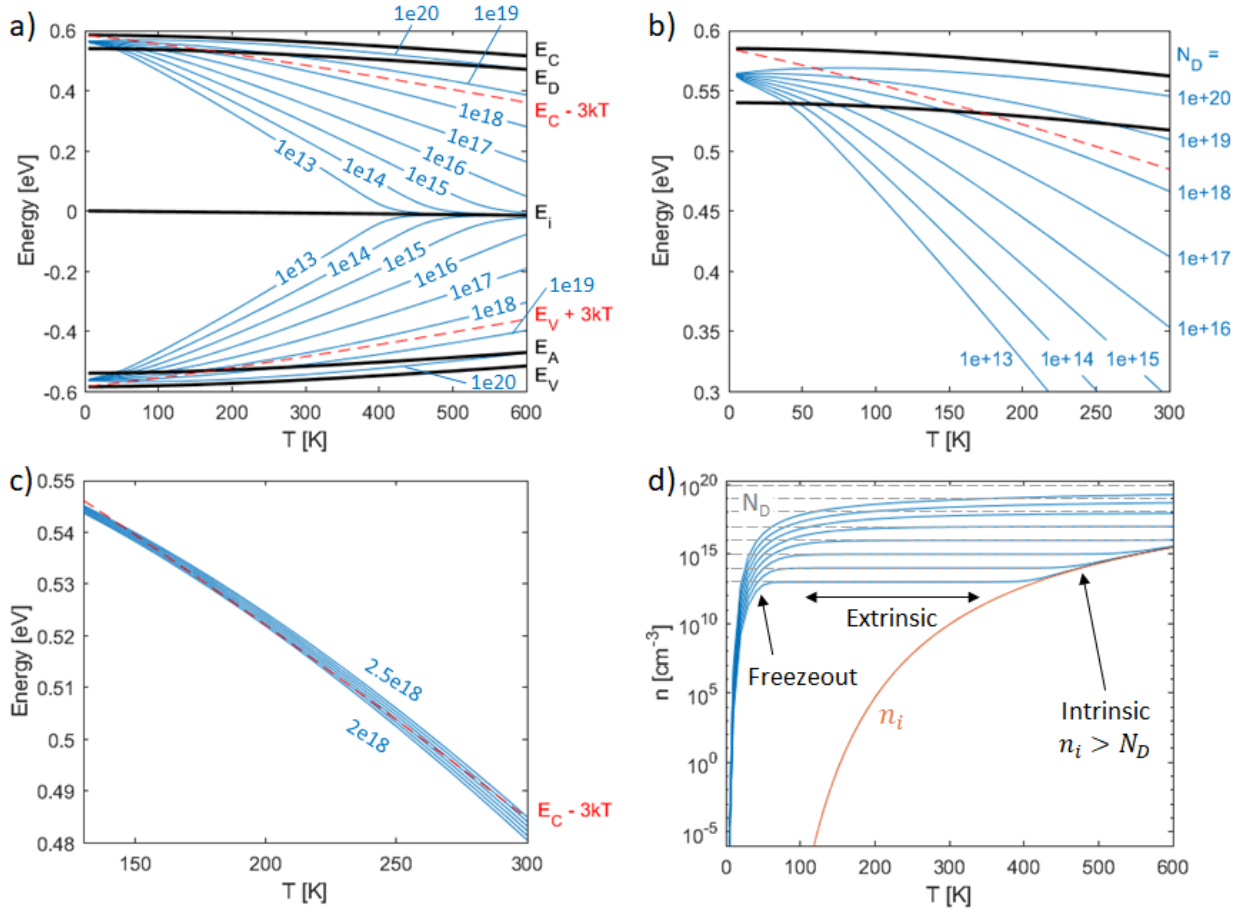


Figure 4.3: Numerical analysis for an Si semiconductor for multiple doping concentrations of Boron (B) or Phosphorus (P). The values for  $g_A$  and  $g_D$  are 4 and 2, respectively. The values used for eq. 4.6 are taken from [70]:  $\alpha = 4.73 \times 10^{-4}$ ,  $\beta = 636$ , and  $E_G(0) = 1.17$ . (a) A wide view of the  $E_F$  vs.  $T$  within the bandgap. The degeneracy threshold (red, dashed line) is defined as  $3k_B T$  off the band gap edges. (b) A zoomed-in view of the  $E_F$  vs.  $T$  relationship in n-type semiconductors at cryoT. (c) A more zoomed-in view into doping concentrations where the semiconductor briefly crosses the degeneracy threshold. (d) The total (solid blue) and intrinsic (solid orange) electron concentrations vs.  $T$ , with  $N_D$  shown in dashed grey lines.

### Semiconductor Ionization Degree

In the equilibrium condition (without electric field nor band-bending), Fig. 4.3d shows the clear transition between the intrinsic ( $n = n_i$  and  $n_i \gg N_D$ ) to extrinsic (where  $n \approx N_D$ ) and freeze-out

( $n \ll N_D$ ). If the doping concentration is too high, the semiconductor may never enter complete ionization ( $n = N_D^+ = N_D$ ), even at RT. Though it may seem related to semiconductor degeneracy, this happens even at  $N_D = 10^{18}$ , a concentration known to be non-degenerate at all  $T$ . Dopants that are not activated (or not ionized) are said to be "frozen-out", while the semiconductor enters the "freeze-out region" when  $n$  significantly drops from  $N_D$ , marked by a sharp kink in Fig. 4.3d. This kink becomes more gradual as the doping concentration increases, and can extend to RT for  $N_D > 10^{19}$ . Therefore, for high doping levels, not only MB approximation is invalid (due to degeneracy), but also the semiconductor is not completely ionized, even at high temperatures. Hence, for simplicity, incomplete ionization must be used for all doping concentrations.

### Freeze-out

While the boundary of the freeze-out region is loosely defined in literature, it may seem reasonable to locate it at the point where  $E_F$  crosses the ionization energy level ( $E_F > E_D$  for n-type, or  $E_F < E_A$  for p-type) [46]. At this point, the incomplete ionization probability (Table 4.1) becomes 33% for n-type ( $g_D = 2$ ) or 20% for p-type ( $g_D = 4$ ). However, high doping concentrations like  $N_D = 10^{20}$  show ionization probability of less than 33% even at 600 K ( $n = N_D^+ < 2 \times 10^{19}$ , Fig. 4.3d), yet the kink in  $n$  occurs at around 90 K. Therefore, to be careful with freeze-out conditions, no special analytical treatment will be applied beyond incomplete ionization.

### Analytical Derivation of the Fermi Level Position

Since the use of MB approximation is validated at cryoT for non-degenerate doping levels, it can now be used to solve the charge neutrality equation 4.3. For simplicity, the equation for n-type and p-type semiconductors is reduced to:

$$N_C e^{(E_F - E_C)/k_B T} = \frac{N_D}{1 + g_D e^{(E_F - E_D)/k_B T}} \quad (4.7a)$$

$$N_V e^{(E_V - E_F)/k_B T} = \frac{N_A}{1 + g_A e^{(E_A - E_F)/k_B T}} \quad (4.7b)$$

Which can be analytically solved to find the relationship between  $E_F$  and  $T$  (see Appendix A):

$$E_F = E_C + k_B T \left[ \ln \left( \frac{1}{2g_D e^{(E_C - E_D)/k_B T}} \right) + \ln \left( \sqrt{1 + 4g_D \frac{N_D}{N_C} e^{(E_C - E_D)/k_B T}} - 1 \right) \right] \quad (4.8a)$$

$$E_F = E_V - k_B T \left[ \ln \left( \frac{1}{2g_A e^{(E_A - E_V)/k_B T}} \right) + \ln \left( \sqrt{1 + 4g_A \frac{N_A}{N_V} e^{(E_A - E_V)/k_B T}} - 1 \right) \right] \quad (4.8b)$$

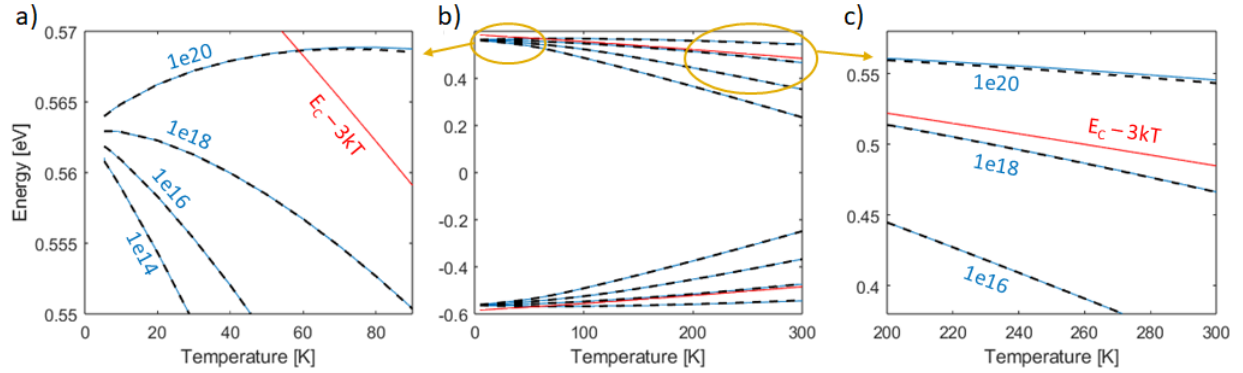


Figure 4.4: Comparison between numerical (solid blue) and analytical (dashed black)  $E_F$  solutions. (b) Shows the entire temperature range while (a) and (c) are zoomed-in views focusing on degenerate doping levels at low and high temperatures, respectively.

To confirm these results, Fig. 4.4 compares them against the exact solution showing an excellent match. It is worth noting that even for degenerate doping levels, eq. 4.8 was only slightly inaccurate above in mid cryoT 4.4c. Moreover, eq. 4.8 is still valid at deep cryoT due to  $E_F$ 's convergence to the midpoint between the band edge and ionization energy.

### 4.1.3 Threshold Voltage

Fig. 4.5a shows the MOSFET device from the vertical perspective (along the y-dimension with the origin set at the semiconductor-oxide interface), focusing on the Metal-Oxide-Semiconductor (MOS) portion of the device. The band diagram for two conditions is illustrated in Figs. 4.5b-c, where applying a potential at the gate terminal induces band-bending in the semiconductor, shifting all energy levels with respect to  $E_F$ .  $E_F$  remains constant because no current is expected to flow through the oxide, maintaining the semiconductor in equilibrium. Internal semiconductor potentials are defined at two main areas of interest (bulk and surface) as follows:

- Internal potential:  $\phi(y) = \frac{1}{q} [(E_i(bulk) - E_i(y))]$   
A measure of the band bending level
- Surface potential:  $\phi_s = \phi(0) = \frac{1}{q} [E_i(bulk) - E_i(0)]$   
Influenced mainly by externally applied  $V_G$
- Fermi potential (also referred to as the bulk potential):  $\phi_F = \phi(bulk) = \frac{1}{q} [E_i(bulk) - E_F]$   
An important material parameter influenced mainly by doping and temperature

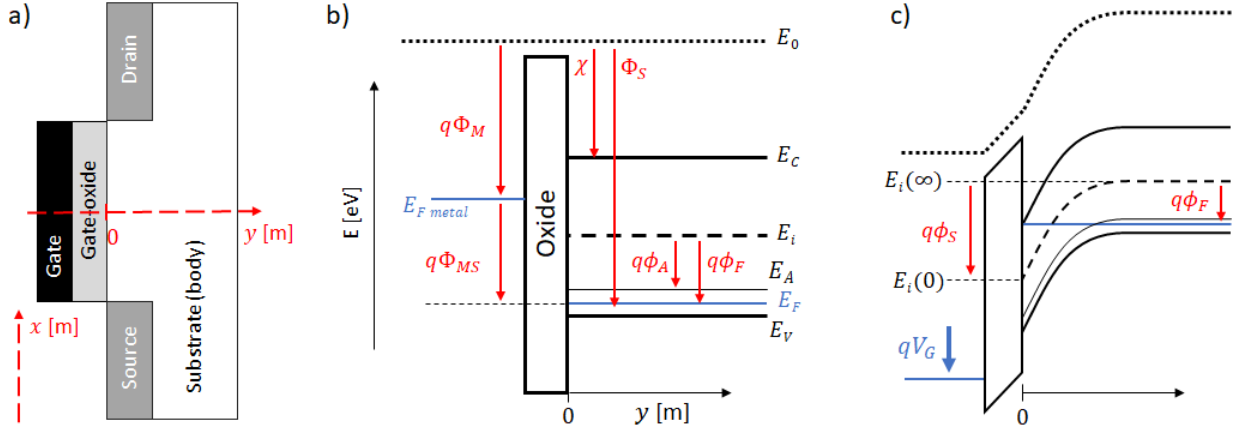


Figure 4.5: (a) An illustration of the the MOS in spacial dimensions focusing on the vertical axis  $y$ . With the origin ( $y = 0$ ) set at the semiconductor-oxide interface (or the surface), (b) and (c) display the band diagram under flat-band and onset-of inversion conditions, respectively.  $E_0$  is the vacuum level,  $\chi$  is the electron affinity, and  $\Phi_M$  and  $\Phi_S$  are the metal and semiconductor work functions who together form  $\Phi_{MS}$ .  $\Phi_S$  can be seen as a combination of  $\chi + \frac{E_G}{2} + \phi_F$ .

- Acceptor potential:  $\phi_A = \frac{1}{q} [E_i - E_A]$   
Constant along  $y$  since it bends with  $E_i$  and is assumed to exhibit the same temperature effect as bandgap widening (maintaining  $E_A - E_V$  constant)

For an NMOS device with a p-type bulk, the relationship between  $V_G$  and  $\phi_s$  is governed by:

$$V_G = \phi_s + \Phi_{MS} - \frac{Q_s(\phi_s)}{C_{ox}} \quad (4.9)$$

where  $\Phi_{MS}$  is the metal-semiconductor work function ( $\Phi_M - \Phi_S$ ), while  $Q_s$  is the charge in the semiconductor (including mobile and localized charge) as a function of  $\phi_s$ . The threshold condition is traditionally defined at the  $V_G$  value where  $\phi_s$  becomes equal to  $2\phi_F$ . At this point,  $V_G$  is referred to as the threshold voltage,  $V_G = V_{th}$ .

Before moving further, there are two frame-of-reference transformations that can provide additional perspective to the analysis in the subsequent subsections. As illustrated in Fig. 4.6, the transformations provide a simple relation between  $E_F$  and  $\phi_s$  (Fig. 4.6b) and make it possible to analyse the effect of varying  $V_G$  from the band bending profile at inversion (Fig. 4.6a)

## Temperature Dependence

To correctly evaluate  $V_{th}(T)$ , one should consider the temperature effect on of each term in eq.4.9.  $\phi_s$  is directly related to temperature (as can be inferred from Figs. 4.3 and 4.6) while  $\Phi_S$  also directly related to  $E_F$  location (Fig. 4.5c). This is not without the assumption that  $E_0$  is constant with temperature, making  $\Phi_M$  temperature independent and allowing  $\chi$  to cancel out the band gap widening effect, due to lack of experimental data in literature [59]. The most relevant charge in  $Q_s$  is the depletion charge  $Q_D$  expressed as [72]:

$$Q_D = -\sqrt{2qN_A\epsilon_{Si}\phi_s} = -Q_{DX}\sqrt{\phi_s} \quad (4.10)$$

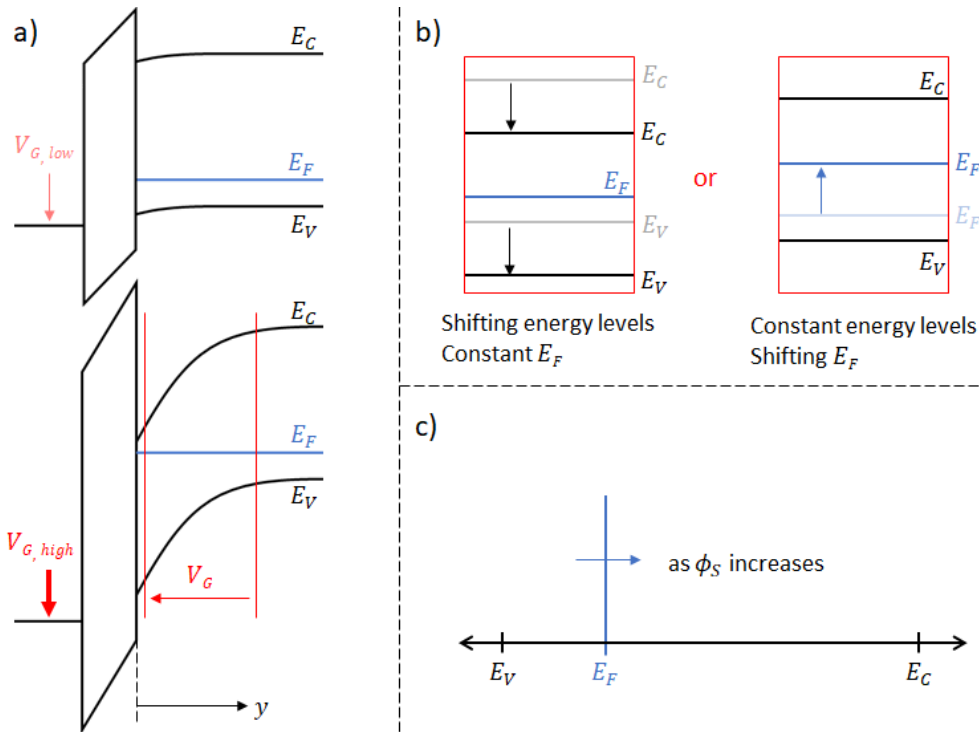


Figure 4.6: (a) Transformation in the frame-of-reference between the spacial dimension  $y$  and the potential dimension  $V_G$ , illustrated as a moving cross-section (red line) tracking the increase in band bending as applied  $V_G$  increases. (b) Transformation in the frame-of-reference where instead of the bands shifting with respect to a constant  $E_F$ ,  $E_F$  shifts within a constant band gap. (c) A 90° view of the band gap, where a change in  $\phi_s$  can be seen as a change in  $E_F$ , or  $\Delta\phi_s = \Delta E_F$ . The remaining energy levels, like  $E_i$  and  $E_A$ , are not shown to improve visibility.

where a  $\sqrt{\phi_s}$  relation with temperature can be derived.  $\epsilon_{Si}$  is the electric permittivity of Si [1] and  $Q_{DX}$  is the temperature independent portion of  $Q_D$ . Finally, the traditional inversion condition must be also revised in cryoT [59]. The relationship ( $\phi_s = 2\phi_F$ ) refers to both areas: the bulk and surface. While the bulk experiences incomplete ionization, the surface exhibits an additional ionization mechanism, i.e. field effect ionization, where the vertical applied field provides the energy required for complete ionization at the surface [24]. Therefore,  $\phi_F$  at the surface would be larger than  $\phi_F$  at the bulk, and the threshold voltage temperature dependence becomes:

$$V_{th}(T) = V_{th}(RT) + \Delta V_{th} \quad (4.11a)$$

$$\Delta V_{th} = \Delta\phi_{FII} + \theta\Delta\left(\sqrt{\phi_{FCI} + \phi_{FII}}\right) \quad (4.11b)$$

where  $\Delta V_{th}$  is the change in  $V_{th}$  due to temperature,  $\theta$  is a fitting parameter  $= Q_{DX}/C_{ox}$ . The subscripts  $II$  and  $CI$  refer to incomplete and complete ionization, respectively. Note that the  $\Delta\phi_{FII}$  from  $\phi_s$  was cancelled out by  $\Delta\Phi_s$ , evaluated at the bulk.

### Analytical Expression for the Fermi Potential

It is apparent from eq.4.11b that obtaining an analytical form for  $\phi_F(T)$  is vital for  $V_{th}(T)$  analysis. The  $T$ -dependence of the Fermi potential in p-type semiconductor can be derived as (Ap-

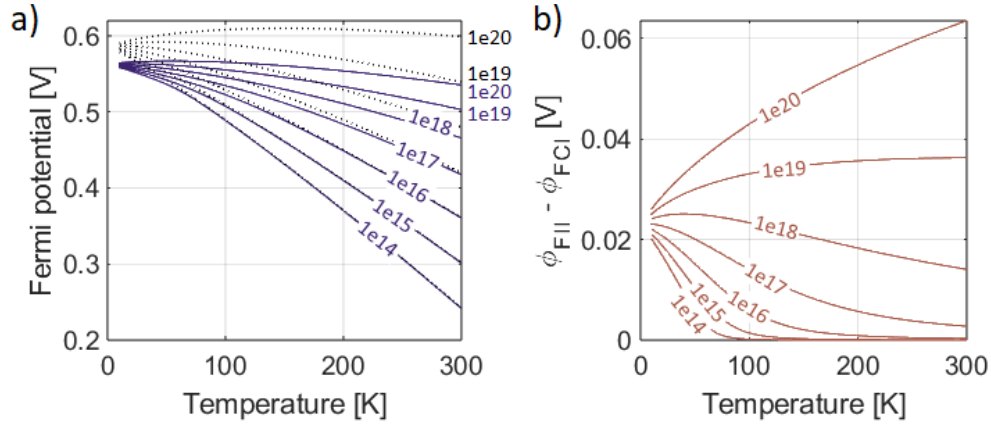


Figure 4.7: (a)  $\phi_{FII}$  (solid) and  $\phi_{FCI}$  (dotted) vs. temperature for multiple doping concentrations. (b) The difference between  $\phi_{FII}$  and  $\phi_{FCI}$  vs. temperature.

pendix B):

$$\phi_{FI}(T) = \phi_T \left[ \ln \left( \frac{N_A}{n_i} \right) - \ln \left( \frac{1 + \sqrt{1 + 4g_A \frac{N_A}{n_i} e^{\frac{-\phi_A}{\phi_T}}}}{2} \right) \right] \quad (4.12)$$

Interestingly, performing the same analysis with while assuming complete ionization yields the simple result (Appendix B):

$$\phi_{FCI}(T) = \phi_T \ln \left( \frac{N_A}{n_i} \right) \quad (4.13)$$

Making the incomplete ionization solution lower than the complete ionization solution by a factor of  $\phi_T \ln \left( \frac{1 + \sqrt{1 + (4g_A N_A / n_i)}}{2} \right)$ . Fig. 4.7 visualizes the difference as temperature decreases for multiple doping levels. Notice the discontinuity in the calculations at temperatures below 10 K. Division by 0 errors are encountered at these low temperatures and will be addressed in section 4.2.2.

#### 4.1.4 Subthreshold Swing

Predictions using eq. 2.1 estimate a vast reduction in the subthreshold swing ( $S$ ) as temperature is reduced, making it one of the most anticipated benefits of cryoT operation for circuit designers. To this day, extensive effort is put forward to produce a cryogenic model suitable for circuit simulations. Since  $E_F$  approaches the band edges at cold temperature, especially during inversion (Figs. 4.3a-b and 4.5c), it was quickly theorized that the elevated interface trap density  $D_{it}$  near the band edges is responsible for the deviations observed in cryogenic  $S$ . Fig. 4.8a[60] shows a typical interface trap distribution within the bandgap, and its exponential increase at energy levels closer to the band edges. However, a full understanding of cryogenic  $S$  phenomena and a complete model of its behaviour is yet to be achieved. Fig. 4.8b-c uses one of the device measurement results in 3.3.2 to illustrate the three cryoT effects on  $S$ : offset, saturation, and  $V_{GS}$ -dependence.

#### Subthreshold Swing Saturation

The  $S$  discrepancy was initially compensated for by a ratio between measured  $m$  at RT and cryoT [21, 22] or an additional temperature-dependent  $m(T)$  term [24, 25, 46, 60]. These models were

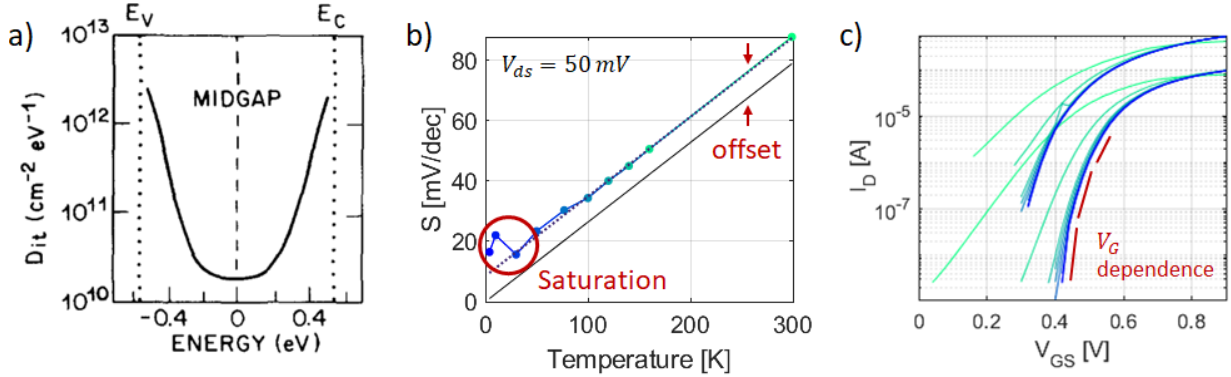


Figure 4.8: (a) A  $90^\circ$  view of the typical interface trap densities within the bandgap at the Si-SiO<sub>2</sub> interface [60]. (b) Measured  $S$ - $T$  profile (markers) showing  $S$  offset and saturation effects at cryoT. The data is shown in comparison to the classical  $S$  theory (eq. 2.1, in solid line) with  $m = 1.33$ , as well as a shifted version by 8.5 mV/dec (dashed line). (c) An illustration of the  $V_G$ -dependence of  $S$  superimposed on the measurement data from Fig. 3.13h.

used in literature until recently challenged by Bohuslavskyi et al. in 2019 [62]. Bohuslavskyi argues that this approach leads to unrealistic  $D_{it}$  values at deep cryoT ( $T = < 4$  K) that exceed the silicon density of states (DoS) of free carriers, and prompted the need for another theory. Band broadening-induced  $S$  saturation was first described in [62] where band tails are placed at the edges of the band gap, blurring the edge of an ideal DoS and affecting the mobile carrier concentration. These band-tails are characterized with a characteristic energy  $W_{tail} = k_B T_S$ , where  $T_S$  is the saturation temperature. The band-tail is attributed to intrinsic mechanisms such as electron-phonon scattering, electron-electron and electron-hole interactions, and finite crystalline periodicity [47, 66]. The  $S$  saturation effect is later modelled using the Gauss hypergeometric function,  $F_1(a, b; c; z)$  [47, 66]:

$$S(T) = m_0 \phi_T \ln(10) \times \frac{\theta^{-1} F_1(1, \theta; \theta + 1; z)}{-z(1 + \theta)^{-1} F_1(2, \theta + 1; \theta + 2; z)} \quad (4.14)$$

where  $\theta = \frac{k_B T}{W_{tail}}$  is unrelated to  $\theta$  in eq. 4.11b, and  $z = -\exp\left(\frac{E_C - E_F}{k_B T}\right)$  at the surface.

### Gate-Voltage Dependence at cryoT

On the other hand, the gate voltage dependence of  $m$  has been first documented in our publication [28] and later investigated further in [61, 62]. It has been attributed to localized disorder-induced

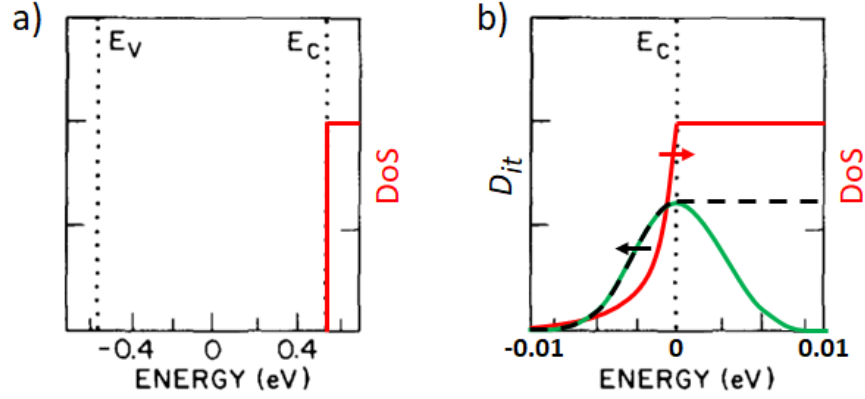


Figure 4.9: Comparison between the traditional view of DoS with sharp band edge (a) and the zoomed-in view showing the band tail proposed by [62] and [61] (b). (b) Also shows the Gaussian distribution of the localized interface traps around  $E_C$  shown in green as proposed in [61], as well as the exponential growth halting at  $E_C$  used in the model discussed in subsection 4.2.2 (dashed black).

interface traps. The most recent model to date is expressed using the Gaussian distribution, chosen to mathematically limit the distribution's growth at and beyond the band edges ( $E_C$  and  $E_V$ ) [61]:

$$Q_o = q \int_{-\infty}^{\infty} G(N_o, W_o) f(E) dE = q \frac{N_o}{2} \left[ \operatorname{erf} \left( \frac{E_F - E_C}{W_o \sqrt{2}} \right) + 1 \right] \quad (4.15)$$

where  $G(N_o, W_o)$  is the Gaussian distribution with magnitude  $N_o$  and width  $W_o$ . However, eq. 4.15 is only valid at deep cryoT where  $f(E)$  can be approximated by a Heaviside step function [61, 73]. Since the effect is exponential in nature, the distribution of these traps can be modelled in different ways as long as it grows exponentially towards the band edges (Fig. 4.9). Therefore, a different solution can be derived using a simpler exponential growth that halts at  $E_C$ . This model is further discussed in subsection 4.2.2.

### Subthreshold Swing Offset

The final cryogenic  $S$  phenomenon is related to uniformly distributed interface traps ( $D_u$ ) that is found to have an offset effect on  $S$  [24]. These traps are already included in the interface traps capacitance ( $C_{it}$ ) found in the traditional threshold factor model:  $m = \frac{\partial V_{GS}}{\partial \phi_s} = 1 + \frac{C_D}{C_{ox}} + \frac{C_{it}}{C_{ox}}$ . However, single-trap analysis reveals their  $\frac{1}{T}$  nature that results from deriving the Fermi-Dirac probability of trap occupation with respect to  $V_{GS}$  [24]. This effect is difficult to detect around

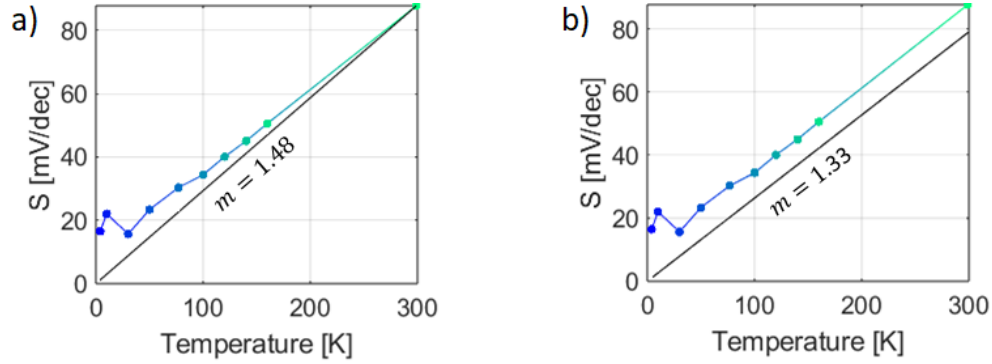


Figure 4.10: Comparison between  $m$  extraction at RT (a) if  $S$  is not taken into account and (b) if it was taken into account. The  $S$  offset in (b) is 8.5 mV/dec.

RT temperatures since it is easily confused with the margin of error associated with experimental measurements. A subthreshold swing of just 1 mV/dec amounts to  $m = 0.017$  at 300 K, but grows to a significant 1.26 at 4 K  $\left(m = \frac{S}{1000\phi_T \times \ln(10)}\right)$ . Fig. 4.10 illustrates this effect on the CMC65n results with an offset of 8.5 mV/dec.

### 4.1.5 Carrier Transport

Carrier transport is well documented in literature to depend on multiple factors that make up the components of both mobility ( $\mu$ ) and velocity ( $v$ ) [15, 29, 32, 35, 63, 64]. In experiments, the on-current is broadly observed to increase with reduced temperature despite the increase in  $V_{th}$  (Fig. 3.12 and 3.13). This is largely attributed to the reduction in phonon scattering with temperature, improving mobility at colder temperatures. The device enjoys this improvement as it approaches deep cryoT up to a point where  $\mu$  starts to plateau or even reverse its trend and starts decreasing. This is because other diffusive scattering mechanisms proportional to  $T$  become small enough to affect the effective mobility ( $\mu_{eff}$ ) and become dominant, resulting in the observed non-and this non-monotonic behavior.  $\mu_{eff}$  combines all diffusive scattering mechanisms according to the Matthiessen rule, where the smallest component dictates the maximum value.

### Reliance on Process and Technology

Coulomb scattering is well-known to be inversely proportional to temperature, making its mobility component proportional to temperature [35]. However, the FET technology and processing

techniques produce other scattering mechanisms related to non-Coulombian defects in the device. These scattering mechanisms are first discovered in [64] where mobility is observed to systematically reduce with channel length. B. Cretu, et al. coined the term "neutral defects" and attributed the effect to crystalline defects near the source and drain regions, as well strain-induced defects. These defects can be healed to a certain degree through temperature anneal during the manufacturing process. Neutral defects are studied further in [63] where it is found that they significantly reduce the mobility's  $T$ -dependence as the channel length is reduced. The effect is attributed to their inverse relation to  $T$ , cancelling the effect of phonon scattering reduction. Other process-dependent effects are also observed throughout literature, complicating the relation between the different mechanisms [15, 29, 32, 35, 74, 75]. With advanced technologies becoming more dependent on layout, even more effects are expected to surface.

### Ballisticity

Other than diffusive scattering mechanisms, there is the ballistic mobility ( $\mu_{bal}$ ) that has a significant thermally driven contribution in short-channel devices (i.e. ballistic and quasi-ballistic devices). It describes the thermionic limit when no backscattering occurs, which is highly dependent on the well-known thermal velocity ( $v_T$ ) and proportionally related to  $T$ . Meanwhile, the velocity  $v$  shows similar non-monotonic behaviour as  $\mu$  and is also governed, in part, by  $v_T$ . Both ballistic parameters combine with their diffusive counter-parts, according to the Matthiessen rule, producing the apparent mobility ( $\mu_{app}$ ) and injection velocity ( $v_{inj}$ ).

### Empirical Model

In addition to ballisticity, carrier transport is heavily dependent on the FET technology (e.g. Bulk, FDSOI, FinFET, etc...), processing parameters (e.g. gate oxide type and thickness, gate length, strain, source/drain formation process, etc..), and how all these components are related to temperature. Therefore, describing the  $T$ -dependence of  $\mu$  and  $v$  analytically will require proprietary information about the fabrication process involved, exhibit different forms for different device technologies, and involve computationally intensive and complex calculations. Instead, an empirical model that encapsulates the general behaviour of all cases is developed for both transport parameters that based on the power-law, a technique frequently used in literature [15, 71, 76–78] (Fig. 2.4b):

$$\mu, v = a_1 T^{p_1} \times WF(T_o) + a_2 T^{p_2} \times (1 - WF(T_o)) \quad (4.16)$$

where  $a_1$  and  $a_2$  describe the magnitude below and above the critical temperature  $T_o$  at which the trend changes, while  $p_1$  and  $p_2$  describe the power-law effect of  $T$  below and above  $T_o$ ,

respectively. Finally,  $WF(T)$  is a window function to facilitate transition between the two trends.

## 4.2 The Compact Model

The physical models developed in previous section may accurately describe the cryoT phenomena observed in measurements, but they are computationally intensive and unsuitable for direct employment in large-scale circuit simulations with multiple devices. Instead, compact models bridge this gap by simplifying the complex physical description into a simple model that is built to be compatible with the numerical methods (like Newton-Raphson method) used by circuit simulators.

The following subsections will introduce the base compact model chosen for this work and its expansion to include cryoT operation.

### 4.2.1 Base Model: MIT Virtual Source Model (MVS)

The MVS model is a simple model for short-channel MOSFETs ( $< 100$  nm) that self-consistently describes the current and charge behavior of nanoscale transistors in the quasi-ballistic regime

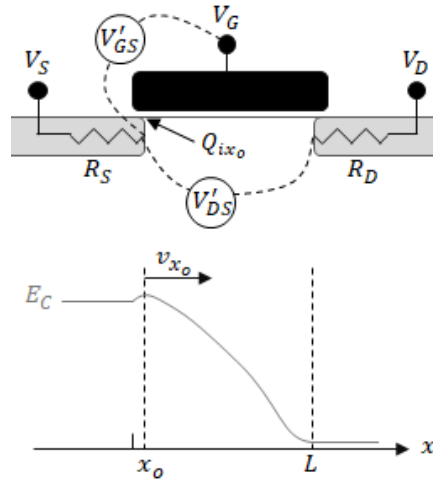


Figure 4.11: Visualization of the MVS model where the VS velocity  $v_{xo}$  and VS charge density  $Q_{ix_o}$  are defined at  $x_o$ .  $V'_{GS}$  and  $V'_{DS}$  are the internal gate-source and drain-source voltages, respectively, excluding the voltage drop across the source/drain series resistances  $R_S$  and  $R_D$ .

[79]. It distinguishes itself from the other compact model by requiring only few fitting parameters (most of which have a physical significance) and being flexible enough to model several FET technologies. The model has been proven to work well for bulk and Silicon-on-Oxide (SOI) MOSFET [79–81], III-V transistors [82] and has also been extended to graphene [83, 84], carbon-nanotube [85], and Gallium-nitride (GaN) transistors [86]. In [87], the basic MVS model has been further upgraded to account for drain-dependent gate charge and carrier density dependent velocity with a few more fitting parameters. These effects are non-trivial in the devices with low carrier effective mass and carrier degeneracy (e.g. III-V devices). However, since the bulk transistors measured in this work are not expected to show significant impacts from these effects, we choose to use the original MVS model as the starting point for simplicity and fewer fitting parameters [88].

$$\frac{I_D}{W} = Q_{ix_o} v_{xo} F_S \quad (4.17a)$$

$$Q_{ix_o} = C_{inv} m \phi_T \ln \left( 1 + e^{\frac{V'_{GS} - (V_{th} - \alpha \phi_T F_f)}{m \phi_T}} \right) \quad (4.17b)$$

$$V_{th} = V_{T0} - \delta V'_{DS} \quad (4.17c)$$

$$F_f = 1 / (1 + e^{[V'_{GS} - (V_{th} - \alpha \phi_T / 2)] / \alpha \phi_T}) \quad (4.17d)$$

$$F_S = \frac{V'_{DS} / V_{DSAT}}{\left[ 1 + (V'_{DS} / V_{DSAT})^\beta \right]^{1/\beta}} \quad (4.17e)$$

$$V_{DSAT} = \phi_T F_f + \frac{v_{xo} L_{eff}}{\mu} (1 - F_f) \quad (4.17f)$$

$$m = m_0 + m_d V'_{DS} \quad (4.17g)$$

In a nutshell, the VS model uses charge-sheet approximation to describe the current as the product of the charge density ( $Q_{ix_o}$ ) and local carrier velocity ( $v_{xo}$ ) at the top of the barrier, or “virtual source” as depicted in Fig. 4.11 and described in eq. 4.17a, where  $W$  is transistor width. In bulk devices where the channel carriers are not degenerated,  $v_{xo}$  is nearly independent of  $V_{DS}$  and  $V_{GS}$  biases.  $Q_{ix_o}$  is approximated by eq. 4.17b,  $C_{inv}$  is the effective gate-to-channel capacitance in strong inversion, and  $V_{th}$  is the threshold voltage after accounting for body effect and Drain-Induced-Barrier-Lowering (DIBL). The Fermi transition function ( $F_f$ ), given by eq. 4.17d, provides smooth transition between subthreshold and above threshold regions using the empirical fitting parameter  $\alpha$ . To describe the current in the linear (or triode) region, the saturation function ( $F_S$ ), given by eq. 4.17e, is employed to smoothly transition between 0 to 1

to signify the transition from the linear to the saturation region with a transitioning factor ( $\beta$ ). This transition is tracked by the  $V_{DS}'$  to saturation voltage ( $V_{DSAT}$ ) ratio.  $V_{DSAT}$  varies between weak and strong inversion as described in eq. 4.17f, where  $L_{eff}$  is the effective channel length. Finally,  $m$  adds the  $V_{DS}'$ -dependent punchthrough factor  $m_d$  to the bias-independent subthreshold factor  $m_0$ .

## 4.2.2 The Cryogenic MVS model

The cryogenic MVS (cryoMVS) model is based on the original MVS model built on the common approximations valid for temperatures around RT and higher. It simplifies the the physical model discussed in section 4.1 for practical use in circuit simulations. Any added cryogenic effects should be compatible with the more advanced versions of the MVS model and easy to integrate. Table 4.3 summarises the changes made to the original MVS model in order to arrive to the cryoMVS model, while the subsequent subsections provide more in-depth discussion of these changes.

### Threshold Voltage

One of the key device parameters used in base MVS model is  $V_{T0}$ , the threshold voltage excluding body effect and DIBL (i.e.  $V_{th}$  when source, body and drain are at the same potential). Although treated as a fitting parameter to allow simplicity and flexibility, theoretically,  $V_{T0}$  should be very close to the gate voltage at the onset of strong inversion. In cryoMVS,  $V_{T0}$  at 300 K is still treated as a fitting parameter, and its temperature dependence is modelled through  $\Delta V_{T0}$  ( $= V_{T0} - V_{T0,RT}$ ). As discussed in section 4.1,  $V_{th}$ 's significant increase at cryoT is dominant by the increase in band bending required to achieve inversion at cryoT, or  $\phi_F$ . This approximation works well to match the measured  $V_{T0}$  trends from across the various process nodes considered.

However the expression developed for  $\phi_F$  in eq. 4.12 possesses a singularity during calculations at very low temperatures. This happens because of the squared exponential relationship with temperature in the intrinsic carrier concentration equation:  $n_i = \sqrt{N_C N_V} e^{-E_G/2k_B T}$ . At very small  $T$ ,  $n_i$  becomes very small (not 0), but interpreted as 0 due to hardware limitations, leading to division by 0 errors. To avoid this, eq. 4.12 is re-arranged to become:

$$\phi_F(T) = \phi_T \left[ \ln \left( \frac{C V_{th}}{T^{1.5}} \right) - \ln \left( 1 + \sqrt{1 + \left( 4 g_A \frac{C V_{th}}{T^{1.5}} e^{\frac{E_{AV}}{k_B T}} \right)} \right) + \ln(2) \right] + \frac{E_G}{2} \quad (4.18)$$

where  $E_{AV} = E_A - E_V$  is a temperature-independent value chosen to be  $\sim 45$  meV for boron implants in the bulk of an NMOS. For the bulk of a PMOS device, a similar quantity of  $E_{CD} \approx$

Table 4.3: Key changes to upgrade the MVS model to the cryoMVS model

| Subject of change  | Original MVS                     | Changes in cryoMVS                                                                                                                                                                       |
|--------------------|----------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Threshold Voltage  | $V_{T0}$ in eq. 4.17c            | Replaced by eq. 4.11 and supplementary eqs. 4.12, 4.13, and 4.18. $c_{V_{th}}$ and $\theta$ are additional parameters. The original parameter, $V_{T0}$ , is preserved as $V_{th}(RT)$ . |
| Subthreshold swing | Eq. 4.17g                        | Replaced by eq. 4.21 and supplementary eqs. 4.19, and 4.20. $T_S$ , $sm$ , $S_0$ , $N_o$ , and $T_o$ are additional parameters.                                                          |
| Mobility           | $\mu$ in eq. 4.17f               | Replaced by eq. 4.23 with additional parameters $a_1$ , $a_2$ , $p_1$ , $p_2$ , $T_{mu}$ , $sm_{mu}$ . The original parameter, $\mu$ , is removed.                                       |
| Velocity           | $v_{xo}$ in eqs. 4.17a and 4.17f | Replaced by eq. 4.24 with additional parameters $b_1$ , $b_2$ , $q_1$ , $q_2$ , $T_{vxo}$ , $sm_{vxo}$ . The original parameter, $v_{xo}$ , is removed.                                  |
| Other              | $\phi_T$ in eq. 4.17d            | Replaced by $\phi_{RT}$ . This applies only to the denominator in the exponential.                                                                                                       |

45meV is used for phosphorus implants. The only fitting parameter introduced in eq. 4.18 is  $c_{V_{th}}$ , which carries the physical meaning of the  $T$ -independent portion of  $N_V$  and  $N_C$  as well as  $N_A$ . Note that the  $T^{1.5}$  factor comes from the  $T$ -dependent portion of  $N_V$  and  $N_C$ . CryoMVS fitting results are shown in Fig. 4.13.

### Subthreshold Swing

The subthreshold swing ( $S$ ) is modelled through the subthreshold factor  $m$  (eq. 4.17g) in base MVS. So far, three cryogenic effects have been identified in cryogenic  $S$  measurements, namely: saturation,  $V_G$ -dependence, and offset (Fig. 4.8b-c). The origins of these effects are summarised in the illustration in Fig. 4.12a where  $S$  saturation is attributed to the blurred band edges of free (or mobile) carrier DoS, while the two latter effects are attributed to the localized (or immobile)

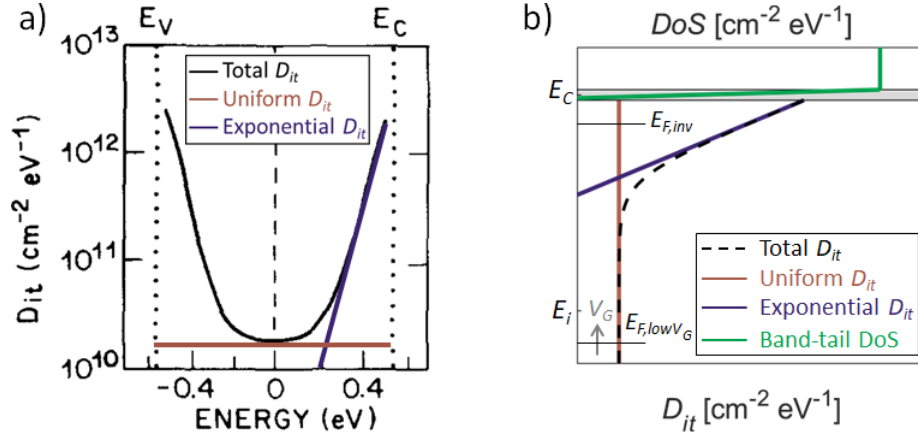


Figure 4.12: Breakdown of  $D_{it}$  as  $D_{ituniform} + D_{itexponential}$ . shown in (a)  $90^\circ$  view [60] and (b) in the band diagram perspective, along with the band-tail DoS.

$D_{it}$  distribution.

The hypergeometric function included in the physical  $S$  saturation model (eq. 4.14) is computationally intensive and contradicts with the efficiency requirement of compact models. Instead, a semi-empirical approach is adopted that also preserves the MVS's flexibility to accommodate various processes with different band-tail profiles. An effective temperature  $T_e$  is defined where it follows  $T$  until the saturation temperature ( $T_S$ ) where it saturates ( $T_e = T_S$  for  $T < T_S$ ). The transition between the two trends is smoothed-out by an empirical dimensionless factor  $sm$  in the following expression:

$$T_e = T_S \left[ 1 + sm \ln \left( 1 + e^{\frac{T - T_S}{sm T_S}} \right) \right] \quad (4.19)$$

The other two major cryoT effects on  $S$  (i.e. offset and  $V_G$ -dependence) are attributed to interface traps summarized in Fig. 4.14a. The figure shows that  $D_{it}$  can be broken into two parts: uniform and exponential distributions. The two effects are additive because they add require interface charge to fill the trap, which increases the amount of applied  $V_{GS}$  required to bend the bands. The uniform portion has the same magnitude at all energy levels within the band gap, and it would result with the same offset no matter where  $E_F$  is located. However, the exponential portion increases in magnitude as  $E_F$  gets closer to the band edge (by applying more  $V_{GS}$  to achieve inversion, Figs. 4.5 and 4.6). Therefore, the  $S$ -offset increases with applied  $V_{GS}$  as it scans exponential  $D_{it}$  levels with increasing magnitudes. The two effects are captured through the subthreshold factor  $m$  in eq. 2.1b as interface charge  $Q_{it}$ . Including interface traps yields two additional subthreshold components:  $m_{S0}$  for  $S$ -offset and  $m_{exp}$  for the  $V_{GS}$ -dependence. These

terms are modeled as follows:

$$m_{it} = m_{S0} + m_{exp} \quad (4.20a)$$

$$m_{S0} = \frac{S_0}{\phi_T \ln(10)} \quad (4.20b)$$

$$m_{exp} = \frac{N_o e^{\frac{V_{gs}^* - V_r}{k_B T_o}}}{\phi_T \ln(10)} \quad (4.20c)$$

$$V_{gs}^* = V_r \left[ 1 - sm^* \ln 1 + e^{-\frac{(V_{GS}' - V_r)}{sm^* V_r}} \right] \quad (4.20d)$$

where  $S_0$  is the  $S$ -offset that accounts for a number of physical parameters including  $D_u$  magnitude, the degeneracy factor of traps, and  $C_{ox}$ .  $V_r$  is a compensated threshold voltage known as the reference voltage ( $V_r = V_{th} - \alpha \phi_T F_f / 2$ ).  $V_{gs}^*$  is a saturated  $V_{GS}'$  with a smoothing factor  $sm^*$  found to be equal to 0.05. Finally,  $N_o$  and  $T_o$  are fitting parameters related to the magnitude and characteristic decay energy of the exponential portion of  $D_{it}$ , respectively. The  $\phi_T^{-1}$  factor originates from the derivative with respect to  $\phi_s$  as discussed in subsection 4.1.4.

Therefore, eq. 4.17g is updated in cryoMVS to become:

$$m(T) = \left( m_0 + V_{DS}' \times m_d \right) \frac{\phi_{T_e}}{\phi_T} + \frac{S_0}{\phi_T \ln(10)} + \frac{N_o e^{\frac{V_{gs}^* - V_r}{k_B T_o}}}{\phi_T \ln(10)} \quad (4.21)$$

changing the subthreshold expression to:

$$S(T) = \left( m_0 + V_{DS}' \times m_d \right) \phi_{T_e} \ln(10) + S_0 + N_o e^{\frac{V_{gs}^* - V_r}{k_B T_o}} \quad (4.22)$$

CryoMVS fitting results are shown in Fig. 4.14.

## Carrier Mobility and Velocity

The carrier mobility and VS-velocity used in the base MVS model represent the physical properties known as the apparent mobility  $\mu_{app}$  and injection velocity  $v_{inj}$ , respectively [87]. These quantities are known to be complex as they are comprised of the combination of their diffusive components (through the Matthiessen rule), which in turn is also combined with the ballistic component (again through the Matthiessen rule). This makes these parameters difficult to predict for a flexible compact model that encompasses a wide range of FET technologies and processes.

Therefore, the general empirical model developed in 4.1.5 is used to model the  $T$ -dependence of cryoMVS's carrier mobility and velocity. This empirical approach allows good flexibility for the users to choose based on the process and device behavior, while still carrying some level of physics. The model equations are formulated as:

$$\mu(T) = a_1 T^{p_1} f(T_{mu}) + a_2 T^{p_2} [1 - f(T_{mu})] \quad (4.23)$$

$$v_{xo}(T) = b_1 T^{q_1} f(T_{vxo}) + b_2 T^{q_2} [1 - f(T_{vxo})] \quad (4.24)$$

where each of the factors (even the critical temperature) are not shared between the two properties. Here, a Fermi function is employed in the temperature dimension instead of energy to act as the window function. It changes from 1 to 0 monotonically with  $T$  to transition between the two trends on either sides of  $T_{mu}$  or  $T_{vxo}$ . CryoMVS fitting results are shown in Fig. 4.15.

## Final Touches

The final adjustment to the MVS model is related to the empirical smoothing function  $F_f$  in the base MVS model (eq. 4.17d). It includes  $T$  in its smoothing factor ( $\alpha\phi_T$ ), where a smaller value means sharper transition. To maintain a consistent smooth transition at all temperatures, it is replaced with the RT value, 300 K.

The cryoMVS is used to describe the  $I$ - $V$  behavior of sample NMOS and PMOS devices from three technology nodes (65-nm, 28-nm, and 20-nm) at cryoT. Fig. 4.16 shows the fitting results for 300K and 4K (in-between temperatures are not shown). These results demonstrate the accuracy of the proposed model to capture  $T$ -dependence of transistor  $I$ - $V$  from RT all the way down to cryo- $T$ . Finally,  $I$ - $V$  characteristics fitting results for 300 K and 4 K are presented in Fig. 4.16.

## Discussion

Fig. 4.17 displays the cryoMVS parameters vs. process node. Note that the x-axes are not necessarily equal to the mask gate length in Table 3.1, and that these samples are not made in the same process. Therefore, one must be careful to not draw assumptions based purely on the gate length, and expect significant influence from the manufacturing process itself.

Fig. 4.17a shows a general trend of increased  $T_S$  with process node advancement, despite the anomaly visible in the PMOS curve. This may be related to the role of finite semiconductor periodicity in blurring the band edges, as pointed out in [66], and the manufacturing trend in

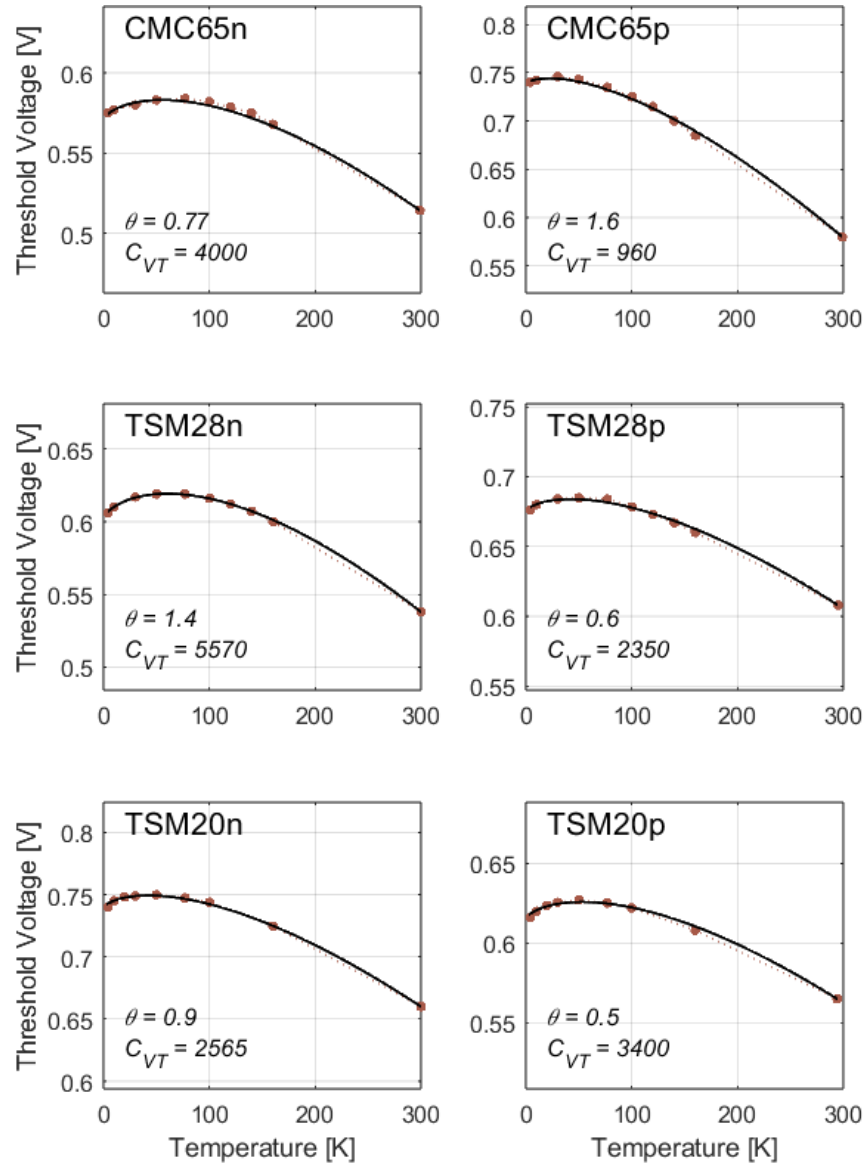


Figure 4.13:  $V_{T0}$  vs.  $T$  of the devices introduced in Table 3.1 showing excellent cryoMVS model fit (solid) against experimental results (markers) for NMOS and PMOS devices from three different process nodes.

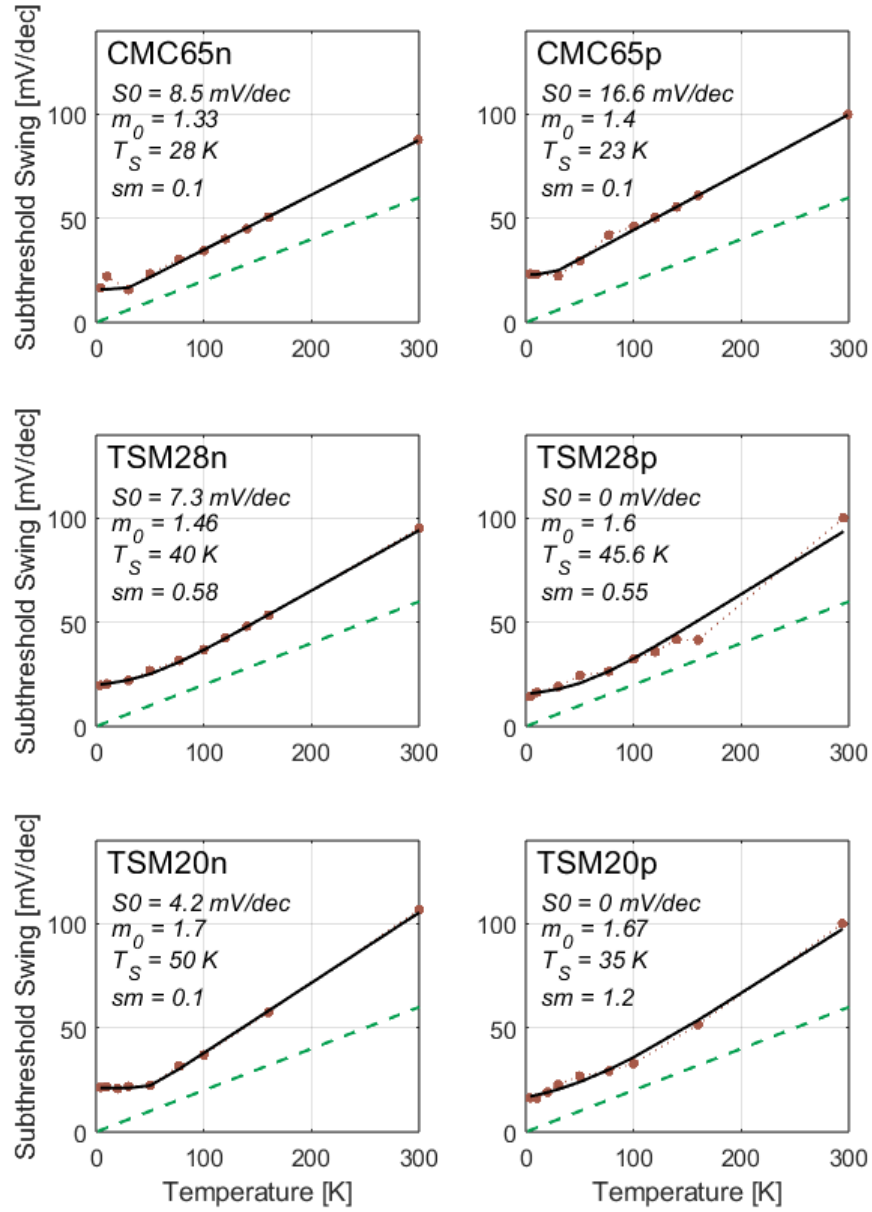


Figure 4.14:  $S$  vs.  $T$  of the devices introduced in Table 3.1 showing excellent cryoMVS model fit (solid) against experimental results (markers) for NMOS and PMOS devices from three different process nodes. The thermionic limit (eq. 2.1a with  $m = 1$ ) is also plotted for reference (dashed).

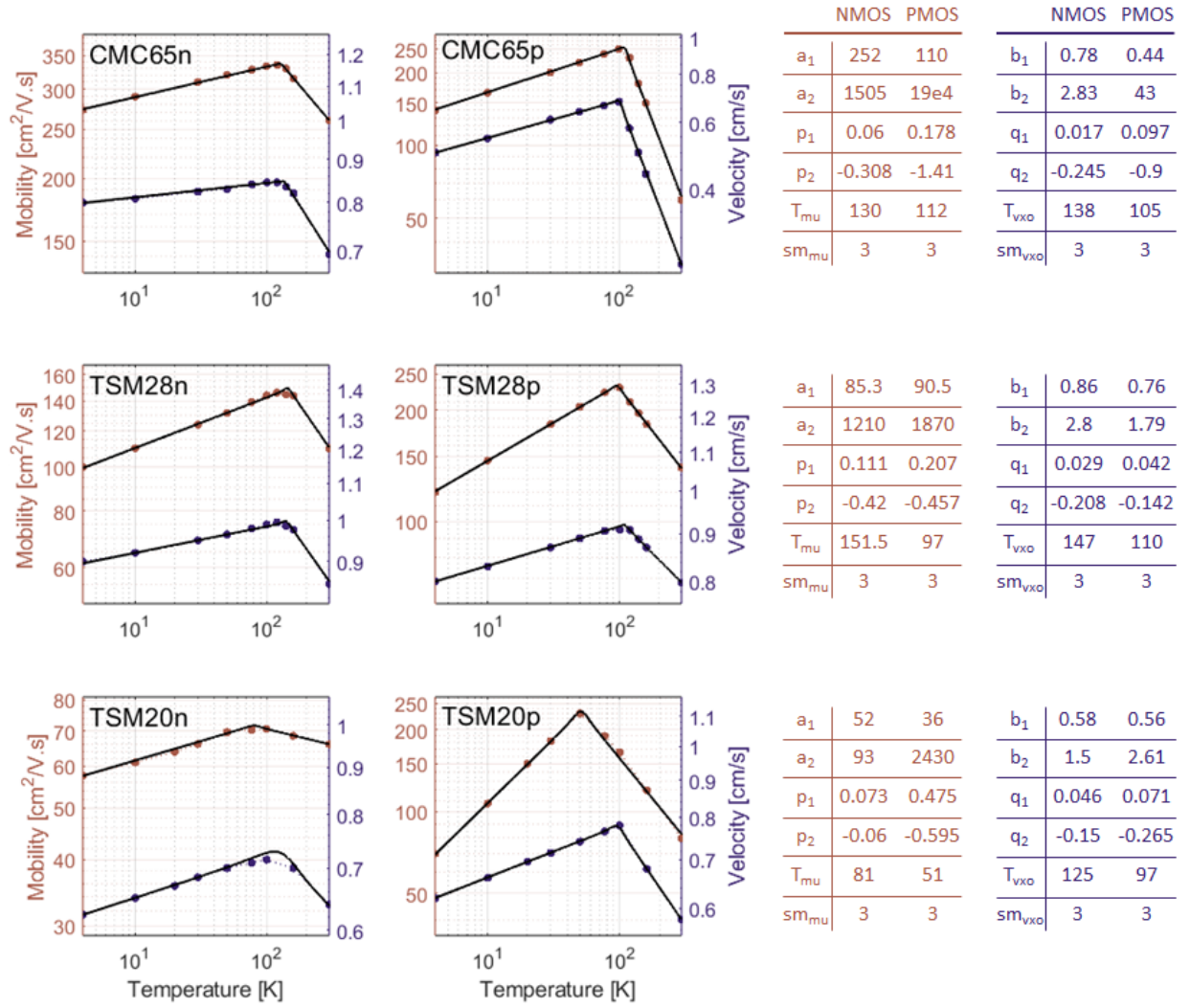


Figure 4.15:  $\mu$  vs.  $T$  (red) and  $v_{xo}$  vs.  $T$  (blue) of the devices introduced in Table 3.1 showing excellent cryoMVS model fit (solid) against experimental results (markers) for NMOS and PMOS devices from three different process nodes.

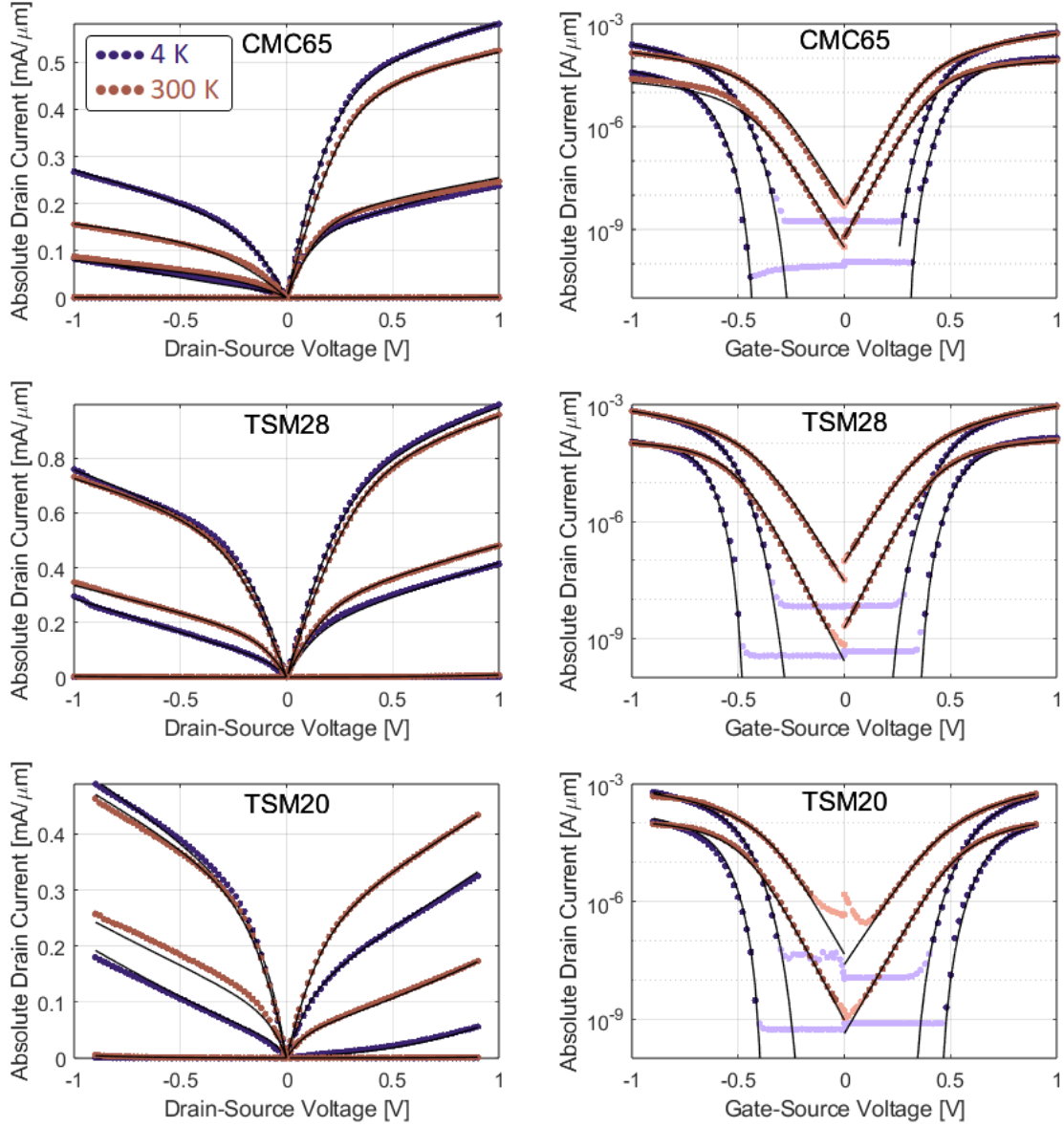


Figure 4.16:  $I$ - $V$  characteristics of the devices introduced in Table 3.1 showing the cryoMVS model fitting results (solid) against measured  $I$ - $V$  (markers). The maximum operating voltage for CMC65 and TSM28 is 1 V, while it is only 0.9 V for TSM20. Output Characteristics:  $V_{GS}$  in CMC65 and TSM28 is set to  $[0.1, 0.7, 1]$  V, while TSM20's  $V_{GS}$  is set to  $= [0.1, 0.6, 0.8]$  V. This is to account for the lower operating voltage in the 20-nm samples. Transfer Characteristics:  $V_{DS}$  is set to  $[0.05, 0.8]$  V. for all devices

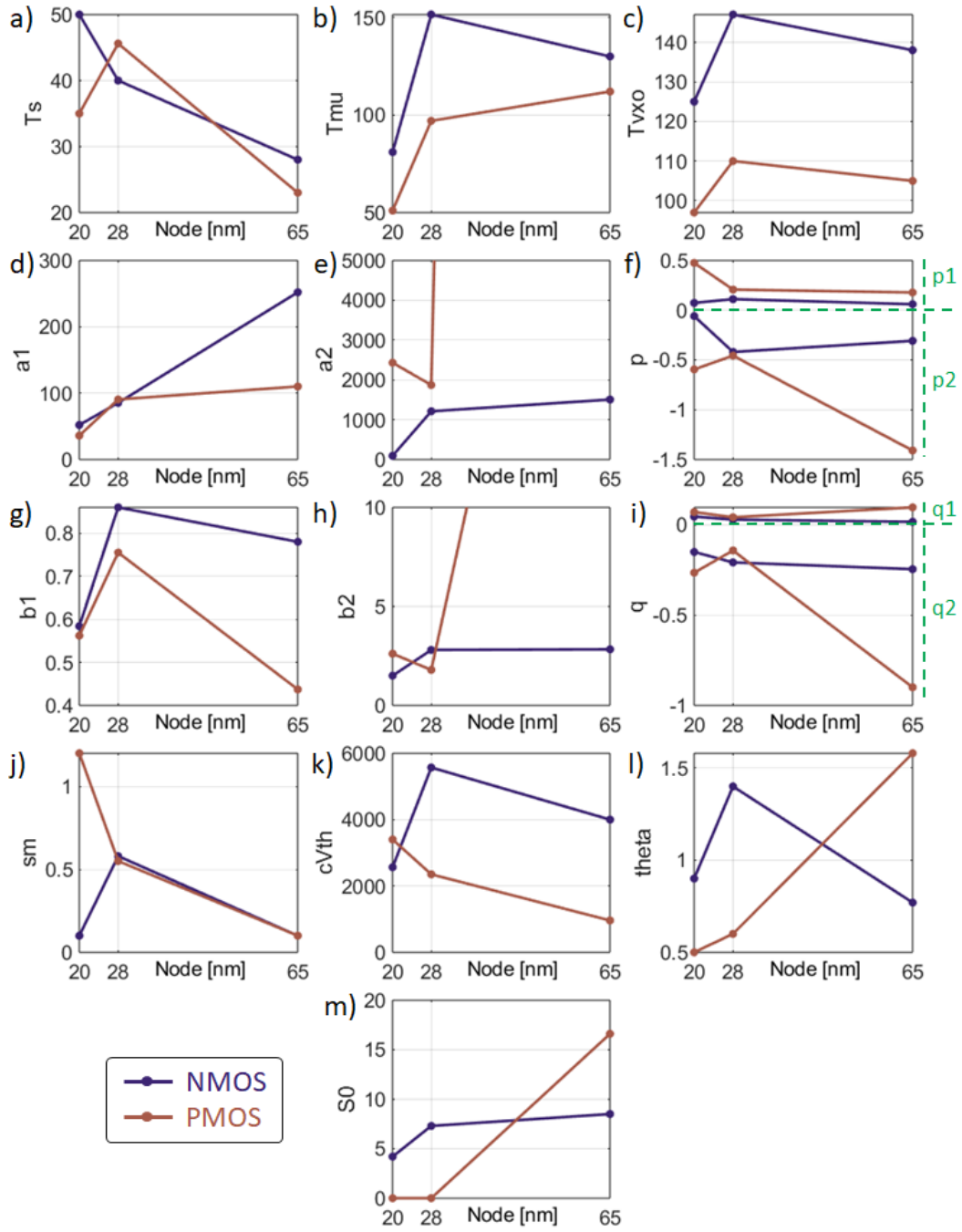


Figure 4.17: CryoMVS parameters vs. process node according to Table 3.1. All parameters were extracted after NMOS and PMOS devices are measured and fitted.

affecting this property. Figs. 4.17b-c do not show a particular trend in  $T_{mu}$  and  $T_{vxo}$  however, they both follow the same behaviour with process node change, suggesting that the same mechanisms affects both mobility and velocity, whether they are gate length related or manufacturing process related. While no device exhibits the same  $T_{mu}$  and  $T_{vxo}$ , a clear relationship between NMOS and PMOS devices stands out where the latter exhibits lower transition temperatures. The available information are not enough to ascertain the cause of this. Moreover, there seems to be no relation between the saturation and transition temperatures ( $T_S$  vs.  $T_{mu}$  or  $T_{vxo}$ ).

The rest of the transport parameters (Figs. 4.17d-i) do not show any particular trend, but a few observations can be made. In general, the magnitude factors of the mobility and velocity ( $a$  and  $b$  parameters) exhibit similar relative values compared to other process node change. For example, CMC65p exhibits very high  $a_2$  and  $b_2$  values compared to other devices (too large to show), while the NMOS  $a_2$  and  $b_2$  curves are very similar. This behaviour is consistent between  $a_1$ - $b_1$ ,  $a_2$ - $b_2$ ,  $p_1$ - $q_1$ , and  $p_2$ - $q_2$ . This is reasonable because the virtual source velocity is expected to be limited by the carrier diffusion at the top of the barrier, as discussed in [87]. However, no general trend can be drawn between NMOS and PMOS. As for the slopes ( $p$  and  $q$  parameters), the expected  $-3/2$  power law in  $\mu$  is not observed in the measured samples, except for CMC65p (most other samples exhibited a slope of around  $-0.5$ ). Moreover, all power law parameters ( $p_1$ - $p_2$ ,  $q_1$ - $q_2$ ) show opposed trends between PMOS and NMOS. When the parameter of one type increases between process nodes, the other type's parameter decreases. Due to the empirical nature of the temperature-dependent transport models, not much can be concluded from these observations, especially with the lack of specific manufacturing knowledge.

$sm$  in Fig. 4.17j is an empirical parameter that models saturation strength that seems to follow the trend of  $c_{V_{th}}$  Fig. 4.17k. A connection can be made through free carrier concentration ( $n$  or  $p$ ):  $S$ -saturation is related to mobile traps, and  $c_{V_{th}}$  contains the effective densities of states in the valance and conduction bands ( $N_V$  and  $N_C$ ), crucial parameters in calculating  $n$  or  $p$ . However, further in-depth investigation is required to be confirmed this sort of connection. With regards to  $S$ -offset ( $S_0$ ), Fig. 4.17m consistently decreases with node advancement. This suggests improved surface passivation with every process node, especially near the mid gap.

Finally, it is difficult to ascertain which trends emerge due to different gate lengths from those dependent on the manufacturing process. Especially with carrier transport, a study of devices with different gate lengths manufactured in the same node may offer more insight. Device ballisticity analysis would also benefit from such a study and greatly improve on inter-node predictions.

# Chapter 5

## Readiness for Practical Simulations

Compact models play a crucial role in circuit simulations. They are the component that describes the behaviour of circuit elements, especially complex ones like transistors, and circuit simulators cannot function without that description. Nevertheless, compact models constitute just one piece of the circuit simulation environment puzzle, making their effectiveness rely heavily on their seamless integration and usability. Therefore, it is imperative that compact models align with circuit simulators and adhere to well-established industry standards to ensure their compatibility and quality. Furthermore, the effectiveness of compact models is diminished if they prove cumbersome to employ, which is undesirable in a rapidly evolving industry. Consequently, technological progress should not be impeded by arduous parameter extraction or time-consuming model fitting.

It becomes crucial to develop complementary tools that work in tandem with the models to overcome these challenges. These supporting tools should streamline parameter extraction processes, alleviating the burden of manual adjustments, and reducing the associated labor and time requirements. Such tools enhance the usability, accuracy, and adaptability of compact models, truly bridging the gap between device physics and circuit simulation, and empowering designers to effectively simulate and optimize complex circuits.

### 5.1 Implementation

To develop a comprehensive and versatile compact model (CM) for device simulation, it is essential to address several key considerations. First, the model should strike a balance between accuracy and efficiency, catering to the needs of both early-stage circuit simulations and

industrial-strength design practices. Device engineers typically use MATLAB for initial model development (such as the case for cryoMVS), a platform renowned for its mathematical capabilities. However, MATLAB's emphasis on mathematical computations falls short when it comes to capturing the behavior of devices within a circuit context. On the other hand, circuit simulators require CMs to be implemented in mid-to-low level programming languages (e.g. C or C++), or more recently, hardware description languages (e.g. VHDL-AMS or Verilog-AMS). These languages contain the necessary properties that integrate well with circuit simulators (e.g. SPICE, Spectre, Eldo, AMS, etc.). However, mid-to-low level languages pose challenges regarding universal operation in different environments.

### **5.1.1 Verilog-A for Compact Models**

A major concern during model development is achieving exact behaviour across the different simulators, operating systems, and their versions. Even the multi-threaded capabilities of each environment would be of concern. Multiple implementations of the same model version can be found for mature compact models, specifically created to address this concern. Moreover, these versions would likely still require re-coding in order to perform properly in current day simulation environments. Another concern is the various optimizations involved in circuit simulation depending on the type of analysis (transient, DC, AC, ...). In C, the developer would be in charge of designing and implementing complex algorithms to accommodate such optimizations. A further challenge is handling communication and information sharing between the CM and circuit simulator, preferably in an efficient manner.

These kinds of burdens should not rest on the CM developer. Instead, the primary focus should be on developing a robust and accurate model that captures the device behavior effectively, while simulation vendors handle uniformity across different platforms, along with the other software details. By shielding the developer from language flexibility issues and prioritizing the model's core functionality and versatility, it becomes feasible to create a single version that can seamlessly integrate with different simulation environments.

### **The Language**

For these reasons, it is decided to implement the cryoMVS model in Verilog-A (VA). The language started as a hardware descriptive language meant for large digital circuits (Verilog), which was then expanded to include digital-analog mixed signal circuits (Verilog-AMS) that contains features inappropriate for compact model use. Verilog-A (analog) is a derivative of the full AMS language that focuses on compact modeling. Much of the interface between the circuit simulator

and the CM is handled by the VA compiler. The compiler also takes care of optimization, as designed by leaders in the industry, and translates the VA code to any simulator on any operating system. The language effectively lowers the bar for becoming a CM developer, as is appropriate for engineers specialized in device physics, not circuit simulators and software development.

Verilog-A took the lead in compact modeling in the early 2000s. It started with 100x performance hit compared to hand-coded C, but became within 2x after a decade. It now receives strong support from EDA vendors like Cadence and Synopsys, and the Compact Model Council (CMC) accepts VA source code for standard model candidates. The industry standard model BSIMSOI transitioned from C to VA in 2010, and BSIM6 is exclusively written in VA in 2013. Other models also use VA nowadays like PSP, Diode\_CMC, EKV, Mextram, and others. Verilog-A has completely been embraced as the standard language for developing and delivering CMs.

## **Best Practices**

Since neophytes to the language don't understand how simulators work or how Verilog-A should work, CMC compiled a set of guidelines and best practices to be followed. The CMC subcommittee is committed to updating those recommendations, and the continuous inclusion of simulation vendors to agree on conventions. There are many ways to describe a certain behaviour in VA code, but only one is agreed upon by the various vendors. Using conventional methods helps compilers recognize certain syntax, which can be used for optimizations like automatic node collapse. This applies not only to new CM developers, but also seasoned veteran ones. For example, one of the classic software best practices from the ASSEMBLY and C era is variable reuse to save memory. This practice is not only fruitless by current-day standards with terabytes of free memory, but adds notable simulation time. If a variable is bias-dependent in one place but not in another, there will be a performance overhead resulting from resetting derivatives for the case where it is not bias dependent, and then bringing them back in when it is bias dependent.

Circuit simulators use modified nodal analysis where node voltages are the unknowns (in additions to certain branch currents) and each row in the matrix specifies a KCL loop for a particular node (or KVL for certain elements). The solution to the matrix is then obtained by Newton's method, which involves function derivatives. Therefore, each node (not model variable) adds an extra system variable, significantly impacting convergence and simulation time. Moreover, the derivative of every mathematical expression in the model must be inspected with respect to any voltage bias for smoothness. any discontinuity and non-smoothness in derivatives with respect to matrix unknowns must be diligently eliminated to ensure convergence.

Finally, all block model capabilities must be avoided such as `transition`, `slew`, `absdelay`, `events`, the `analysis` block, etc... Only the analog block is required for compact modeling.

These capabilities are left available in the language despite being problematic for CMs because VA is used in analog applications other than compact modeling.

### 5.1.2 MVS model Verilog-A implementation

Being essentially a software, CM development would greatly benefit from good software practices, including version control. For this reason, an intermediary version of the base MVS model is created, where all the best practices are ensured to be followed, before developing it into the cryoMVS model.

#### Improving the Base Model's Verilog-A Implementation

Device engineers often use centimeter (cm) as a unit of dimension, but circuit simulation environments conventionally use the international system of units (SI) version, which is just meter (m). The base MVS is implemented from device engineering perspective, which is problematic because circuit designers will enter values based on the SI units, producing undesired results. Therefore, the units of the model are corrected to comply with SI units.

Moreover, it will be difficult to adjust the vast number of circuit design automation tools to provide special treatment to the MVS model, defeating the purpose of having a CM implementation. Moreover, CMC requires CMs to define the acceptable range of every input parameter to prevent mathematical errors, without need for smoothing or absolute value conversion. For example, the `phib` parameter in the MVS model is left without a range. When `phib` appears in square root:

```
Vt0bs0 = Vt0 + gamma * (sqrt( abs( phib - Vbsi )) - sqrt( phib ));
```

knowledgeable users might know that it must be negative, but automated software and unsuspecting users can input negative values, leading to complex numbers and errors. The same issue was also present for `Vt0` and `alpha`.

Another common shortcoming found in the VA implementation is the syntax used to access voltage values:

```
Qsov = Cofs * ( V(g) - V(si) );
```

This syntax retrieves the potential between the specified node and global ground. This may not be an issue in this particular case since both nodes are with respect to the same ground, but CMC requires the use of branch voltage syntax to obtain potential difference between the two nodes directly. Therefore, the correct implementation would be:

```
Qsov = Cofs * V(g,si);
```

Finally, CMC requires proper coding style to ensure maintainability and transferability, which was violated on multiple occasions. This also includes indentation using spaces, not tabs, to ensure uniform presentation on all text editor software.

## CryoMVS Verilog-A Implementation

There are several important guidelines that need to be upheld during the conversion of the cryoT part of the cryoMVS model. First, the best accepted physics constants (like  $q$ ,  $k_B$ , etc...) are continuously updated. A fitted model based on an old set may behave differently on an updated environment, making these constants unintentional parameters of the model. To ensure consistency the recommended practice is to hard-code these values into the model, [89]. Moreover, double digit arithmetic must provide enough significant digits for accurate calculations, about 17 digits:

```
'define PI = 3.14; // not recommended
'define PI = 3.14159265358979323; // has enough significant digits
```

Common arithmetic constants like  $\sqrt{2}$  and  $\log_{10}(e)$  are provided in the Verilog-A compact modeling extension constants.vams. In addition, a few more have been contributed like  $\sqrt{\pi}$ .

A common mistake associated with converting CMs from MATLAB is that the syntax for natural logarithm in MATLAB,  $\log(x)$ , is interpreted as  $\log_{10}(x)$  in VA. Thus, these functions should be replaced with  $\ln(x)$ . It is also important to be mindful of expensive arithmetic operations such as: exponential, power, logarithm, etc... Using predefined constants like  $\log(2)$  and  $\sqrt{0.5}$  helps reduce simulation time, but piecewise definition of complex equations is much more effective. For example, eq. 4.20d contains the computationally expensive  $\exp$  and  $\ln$  functions. Since it is essentially a smoothing function between  $V_{GS}$  and  $V_{th}$ , it can be optimized as:

```
etastar = (Vgs - Vth) / sm / Vth;
if (etastar <= -6) begin
    Vgstar = Vgs;
end else if (etastar >= 6) begin
    Vgstar = Vth;
end else begin
    Vgstar = Vth * (1 - sm * ln( 1 + exp( -etastar ) ));
end
```

Making the computationally costly calculation happen only when needed (Fig. 5.1)

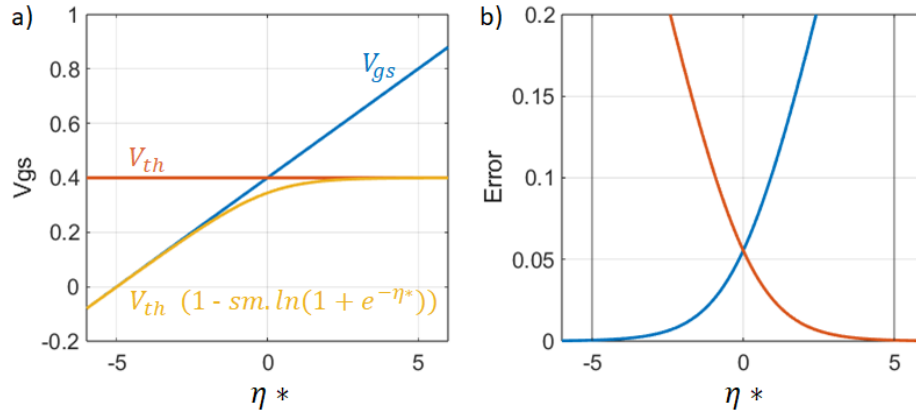


Figure 5.1: Illustration of piecewise optimization, a useful technique for CM models. (a) A comparison between the full form function and the left and right approximations. (b) The error (the difference) between the full form and approximations.

It is also paramount to be mindful of derivatives, especially when piecewise equations.

Other common guidelines include:

- Use meaningful variable names:

$v_{ds}$  instead of  $v_a$  or  $v_x$

- Do not rely on character case as SPICE is case insensitive:

$VAR$  is the same as  $var$  and  $vAr$

- Add decimal points to constants to avoid integer division:

$1/2 = 0$  but  $1.0/2.0 = 0.5$

- Capacitor current must be obtained from the derivative of charge, not voltage. This is to prevent the accidental loss of charge conservation if the capacitance is bias-dependent, and follow the agreed upon convention:

$$I = \frac{\partial Q}{\partial t}, \text{ not } C \frac{\partial V}{\partial t}$$

- Avoid hidden states in `if` and `case` statements by making sure that relevant variables are defined for all cases. In the following example, a hidden state for `h` is used when `dim = 2`:

```

if (dim == 3) begin
    h = ...
    w = ...
    l = ...
end else begin
    w = ...
    l = ...
end
end

```

## Testing

Like any piece of software, compact models must be routinely tested during the development phase and prior to release. Testing the model within the intended region of operation is a great way to ensure correct behaviour, as unexpected voltages far from the realistic range could be applied during the circuit matrix solving process. The model must be tested at such extreme voltages for graceful handling without any mathematical errors. Correctness in this case is not needed since the converged values will be within the intended operation range.

Since MVS is written for symmetric FET devices, this symmetry must be also tested while sweeping across the  $V = 0$  mark and other symmetry points, for both n-type and p-type. There is a symmetry standard called the Gummel symmetry test where a differential voltage  $2V_x$  is applied at the source and drain terminals as shown in Fig. 5.2a. To ensure symmetry, the current  $I_x = \frac{I_D - I_S}{2}$  is examined for any discontinuities around the symmetry point (in this case:  $V_{DS} = 0$ ). The Gummel test can also examine the model's readiness for RF applications where multiple orders of  $I_x$  derivatives are also monitored for smoothness. Since the cryoMVS model is developed in the DC perspective, only the first two derivatives are investigated in Figs. 5.2c-d. For an RF model, the smoothness of more derivatives must be validated (up to six or more). In this example, a discontinuity is visible in the  $V_G = 0.4$  V curve (Figs. 5.2d), indicating that cryoMVS is only ready for static and light transient simulations.

Finally, the model must be tested for consistency in the circuit simulation engines available on The University of Waterloo's EDA servers: HSPICE and Spectre. Spectre is known to be very susceptible to hidden states, while HSPICE is more commonly used for DC applications [89]. Both simulators provided identical results despite their differences, verifying the effectiveness of the CMC guidelines. The VA code is provided in Appendix C.

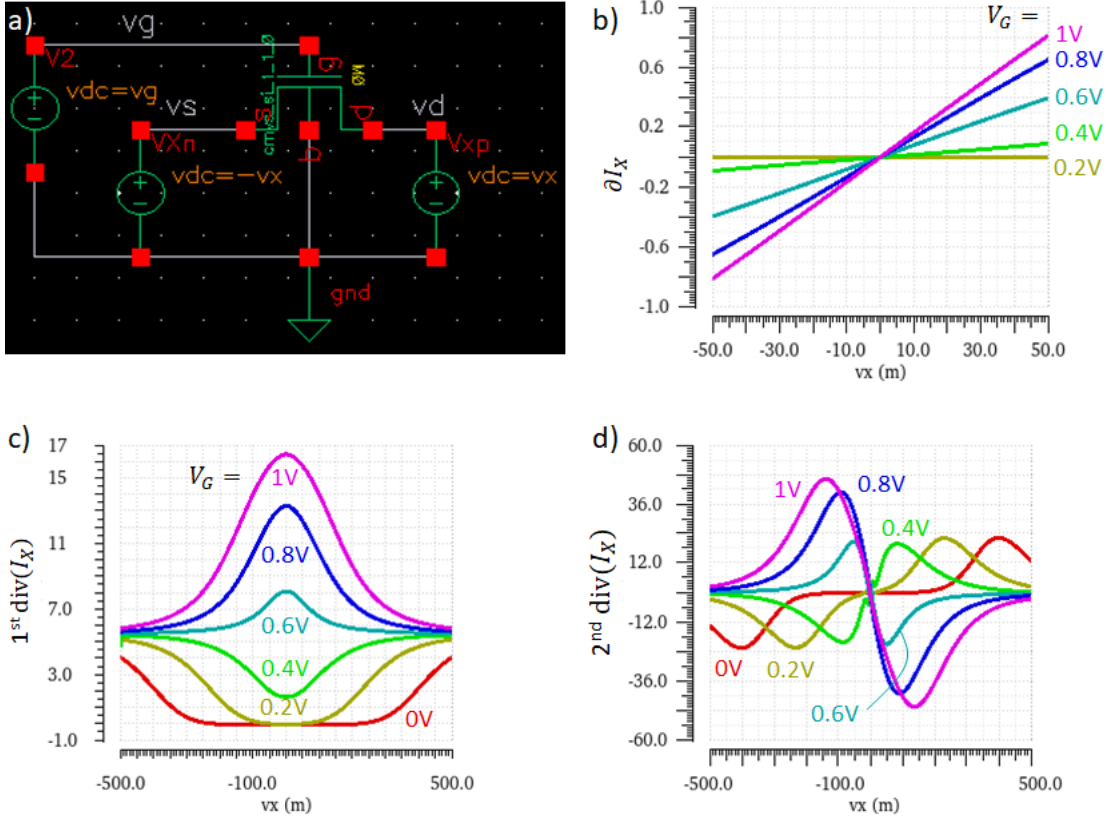


Figure 5.2: (a) Circuit setup for the Gummel symmetry test. (b)  $I_x$  continuity around the symmetry point  $V_{GS} = 0$ . (c) The first and (d) the second derivatives of  $I_x$  depicting the model's smoothness around the symmetry point. The slight discontinuity observed in the  $V_G = 0.4$  V curve of the second derivative is not a major concern for this DC model, but a sign that a dynamic cryo-model is necessary for intensive analog or RF simulations.

## 5.2 Supporting tools

Once the implementation of a compact model is complete, CMC emphasises that the VA code alone is not enough. The accuracy of the model parameters utilized directly impacts its fidelity and predictive capabilities. Therefore, the next crucial step involves parameter extraction and model fitting, which prepares the model to accurately describe devices from a specific process node and facilitates circuit simulations. This process typically involves measurements from real devices (or detailed physical simulations) under specific conditions, best known by the CM developer, and then fitting the model equations to align with the experimental data. Given that these

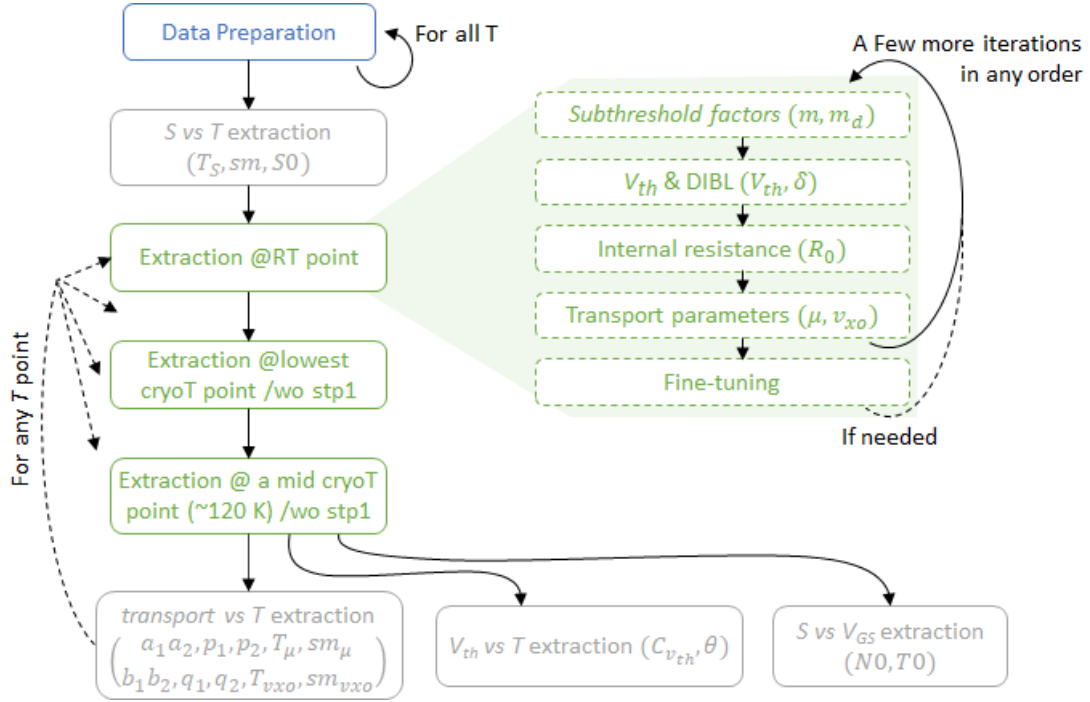


Figure 5.3: Diagram for recommended extraction flow to maximize parameter extraction and model fitting efficiency. There is also an option available to auto-fit  $I$ - $V$  characteristics for all temperature points (not included in the flowchart). This recommended flow is not to be taken as a rigid guideline. Other extractions flows may be more convenient depending on the specific DUT and measurements made.

tasks are typically carried out by technology vendors rather than the CM developers themselves, it is important to include proper documentation and working sample parameter set, along with its matching experimental data.

Fig. 5.3 presents the general flow of cryoMVS parameter extraction. Each step represents a separate tool created to facilitate and streamline the parameter extraction and model fitting procedures. These tools enable technology engineers to efficiently calibrate the model, and together with automation routines, reduce manual effort and mitigate potential errors associated with manual tuning. As a result, reliable and consistent results throughout the process are ensured. Moreover, streamlining the workflow is essential for the ongoing development and refinement of CMs. As new technologies continue to emerge and device characteristics evolve, these tools enable researchers and engineers to efficiently adapt and update their models, allowing them to remain relevant and accurate.

## 5.2.1 Parameters Extraction

A MATLAB software is developed to streamline measurement data processing, parameter extraction, and model fitting. The software is written from an object-oriented perspective along with graphical user interface (GUI) for speedy results.

A typical data set is expected to have the one-measurement-per-temperature format where a lot of unnecessary output and transfer curves are present (Fig. 3.11). DC analysis is commonly conducted on a few output and transfer curves (Figs. 3.12, 3.13), and hence, there is an option to choose the  $V_{DS}$  and  $V_{GS}$  biases desired. Before parameter extraction, some data preparation might be necessary. It is worth noting that applied  $V_{GS}$  and  $V_{DS}$  biases are used for output and transfer characteristics, respectively.

### Data Preparation

As mentioned in section 3.3.2, both output and transfer characteristics are measured simultaneously, generating measurement data that require post-processing before viewing and examining. The first supporting tool performs the task of rearranging the data points for  $I_D$ - $V_D$  and  $I_D$ - $V_G$  perspectives, while selecting specific curves based on the user's preference (Fig. 5.4)

Next, identifying anomalous data points that are not modeled is an essential step prior to utilizing automated parameter extraction and fitting routines. Such data points are often found in the low  $V_{GS}$  range of the transfer characteristics and arise from various leakage mechanisms such as gate-induced drain leakage (GIDL). However, the main source of anomalous data points in this regime is the minimum compliance current of the measurement instrument. Though this may not be a significant problem at room temperature, at cryogenic temperatures, the subthreshold current levels are orders of magnitude smaller and they quickly drop to a point where they become smaller than the minimum current that can be detected by the instrument. Fig. 5.5 shows the tool at which these data points are identified utilizing an example at 4K. Note that no data points are deleted, but only identified to be excluded during automated extraction.

At this point, the designer is ready for the parameter extraction flow which consists of several steps that may require multiple iterations to produce a decent fit for a DUT. There are several correct ways to traverse this process, but the flowchart in Fig. 5.3 presents the recommended flow to avoid common mistakes and efficiently obtain good fits with just few iterations. For example, some  $T$ -dependence extraction steps (like steps 7 and 8 cannot be used until fitting parameters for at least three temperatures are available.

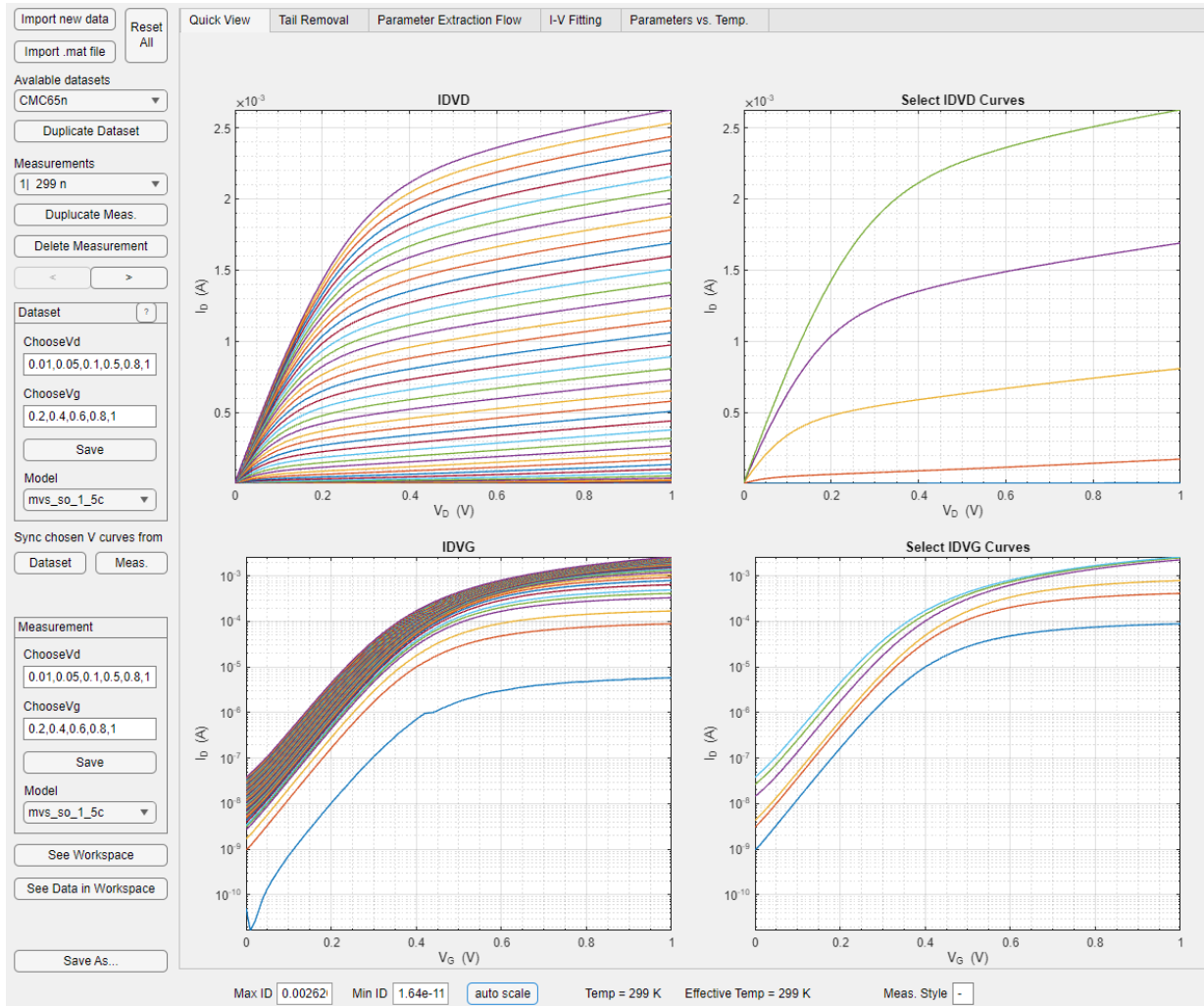


Figure 5.4: Quick View tab of the MATLAB application showing speedy post-processing of raw data along with a number of convenient options.

## 5.2.2 Single Temperature Extraction

The extraction is done at one temperature point at a time, with the RT point first, then the coldest point before the mid temperature points. This is to obtain a general idea of the parameter range that will guide the designer through most points. Fitting at any temperature point can use the parameter set from another temperature point as a starting guess, while monitoring the parameter change trend.

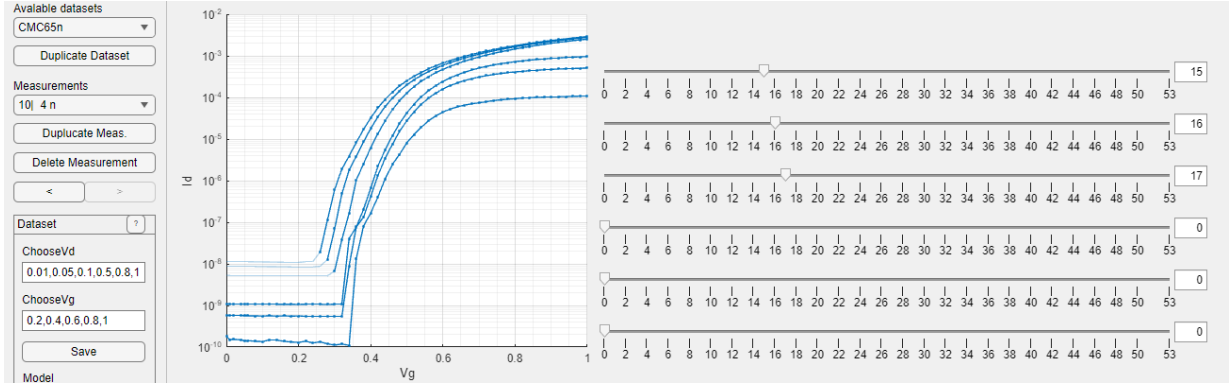


Figure 5.5: Data preparation tool to remove obstructive data-points from transfer characteristics.

The Extraction process in the base MVS model involves curve fitting  $I_D$ - $V_D$  then  $I_D$ - $V_G$  sequentially in a loop, obtaining a rough estimate at first and then a more closer parameter set for the device in question. This method of fitting all MVS parameters at once will be utilized to fine tune the parameter set in section 5.2.2. The following extraction steps eliminates the effects of temperature to allow focusing on device parameters in a given temperature point based on practical measured data. Though the result is a parameter set per temperature, temperature-related fitting parameters can then be swiftly extracted in section 5.2.3 to produce only one parameter set per device. For each step, the tool will aid the user by provide visual guides (presented as black dashed line) as well as automated extraction where possible. The tool will also show model reconstruction of the  $I - V$  curves in view as a black dotted line, which will gradually falls into place to fit the data-points as more parameters are extracted.

### Step1: The Subthreshold Factor (Related Parameters: $m_0$ , $m_d$ )

This step will focus on the extraction of  $m$  in the linear region, usually performed on the  $V_{DS,lin} = 50$  mV transfer curve. By default, the tool will chooses this value, or if not available, the lowest  $V_D$  bias available. Note that the extraction is based on the classical subthreshold equation 2.1a without the context of effective temperature nor  $S$ -offset. The factor  $m$  can be calculated from measurement data as:

$$m = \frac{S}{\phi_T \ln(10)} \quad (5.1)$$

The visual aid originating from the lowest  $(V_D, I_D)$  data-point with a logarithmic slope of  $m$ . In case the user wants to use a different  $(V_D, I_D)$  point, the option to change the origin point is available through adjusting ID0 and VG0 (Fig. 5.6). Meanwhile, the developed auto extraction

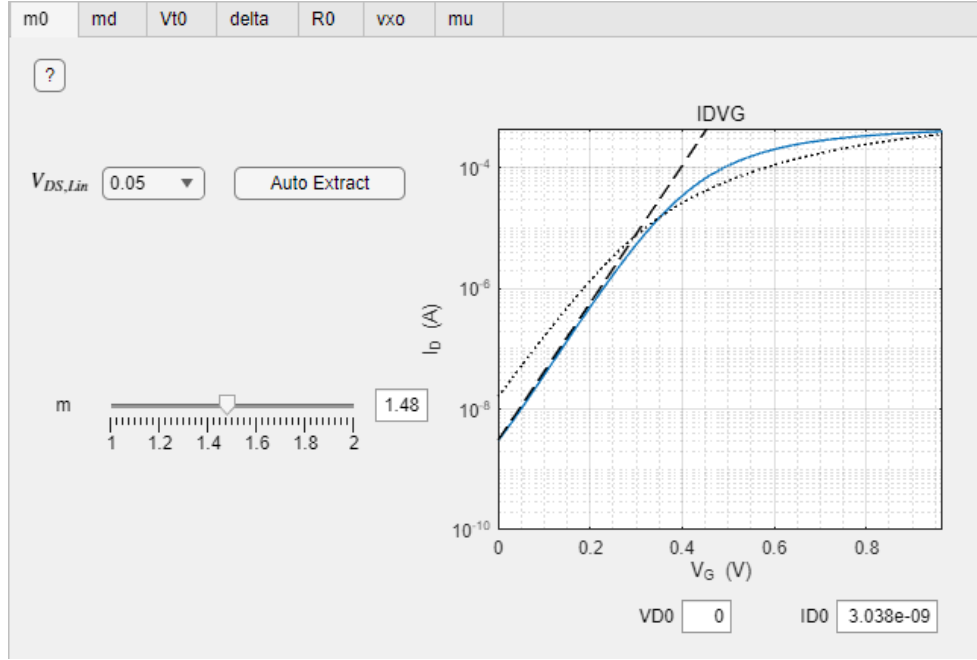


Figure 5.6:  $m$  extraction GUI for a single temperature.

routine is based on eq. 5.1 using the transfer curve associate with the chosen  $V_{DS,lin}$  bias. The routine uses the smallest logarithmic slope in the curve, which is usually the first two ( $V_D, I_D$ ) points in the transfer curve. This should serve as an initial guess, and its accuracy varies based on the provided data's resolutions and amount of noise present.

The extracted  $m$  serves as a good starting point for the actual parameters used in the model:  $m_0$  and  $m_d$ . Similar to the DIBL effect on  $V_{th}$ ,  $m_d$  models the  $V'_{DS}$  on  $m$  known as punchthrough effect, while  $m_0$  is devoid of any punchthrough effect ( $V'_{DS} = 0$ ). Indeed, the previously extracted  $m$  contains the effect of  $V_{DS,lin}$  bias. Punchthrough is modeled as a linear effect with  $m_d$  as a constant fitting parameter (eq. 4.17g). The extraction can be performed on  $m_0$  and  $m_d$  pair or the  $m_{Lin} = m_0 + m_d V_{DS,lin}$  and  $m_{Sat} = m_0 + m_d V_{DS,sat}$  pair. The relation between the two pairs can be expressed by the following linear system:

$$\begin{bmatrix} m_{Lin} \\ m_{Sat} \end{bmatrix} = \begin{bmatrix} 1 & V_{DS,lin} \\ 1 & V_{DS,sat} \end{bmatrix} \cdot \begin{bmatrix} m_0 \\ m_d \end{bmatrix} \quad (5.2)$$

A similar extraction is employed, where a saturation transfer curve is added, usually at  $V_{DSAT} = 0.8$  V. The visual aids and automation routine in this step are based on the  $m_{Sat}$  and  $m_{Lin}$  pair.

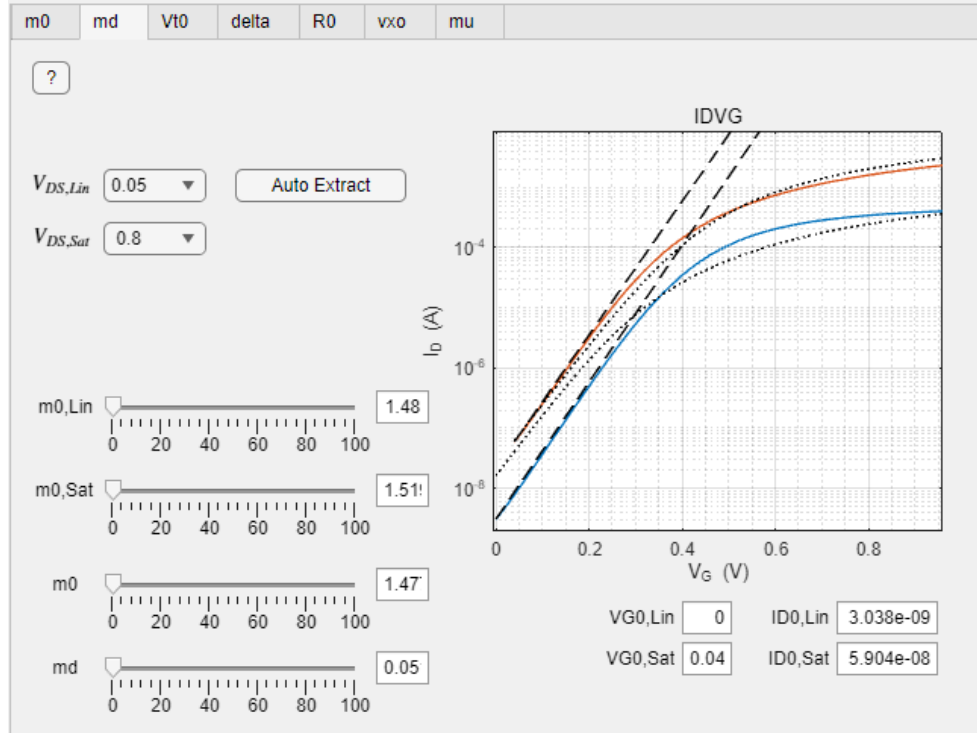


Figure 5.7:  $m_0$  and  $m_d$  extraction GUI for a single temperature.

The user can manually adjust either pair and the tool will automatically deduce the values for the other pair according to eq. 5.2 (Fig. 5.7).

## Step 2: Threshold Voltage (Related Parameters: $V_{T0}$ , $\delta$ )

Threshold voltage is a more subjective device parameter than would be expected. While a concrete physical definition exists (eq. 4.9), it relies on physical parameters that difficult to measure directly (e.g.  $\phi_F$  and  $\phi_S$ ). Hence,  $V_{th}$  is extracted from measurement data indirectly using various methods [90], leading to different measurement-based definitions.

Since the MVS model is not surface-potential-based and a simple model, it treats  $V_{th}$  as a fitting parameter. Therefore, a good estimate can be achieved visually by locating the kink between the subthreshold and above threshold regimes in the linear transfer curve. This is performed by using the slider to position the visual aid (dashed line) within the kink in the transfer curve (Fig. 5.8), usually taken at  $V_{DS,lin} = 50$  mV with an option to change it.

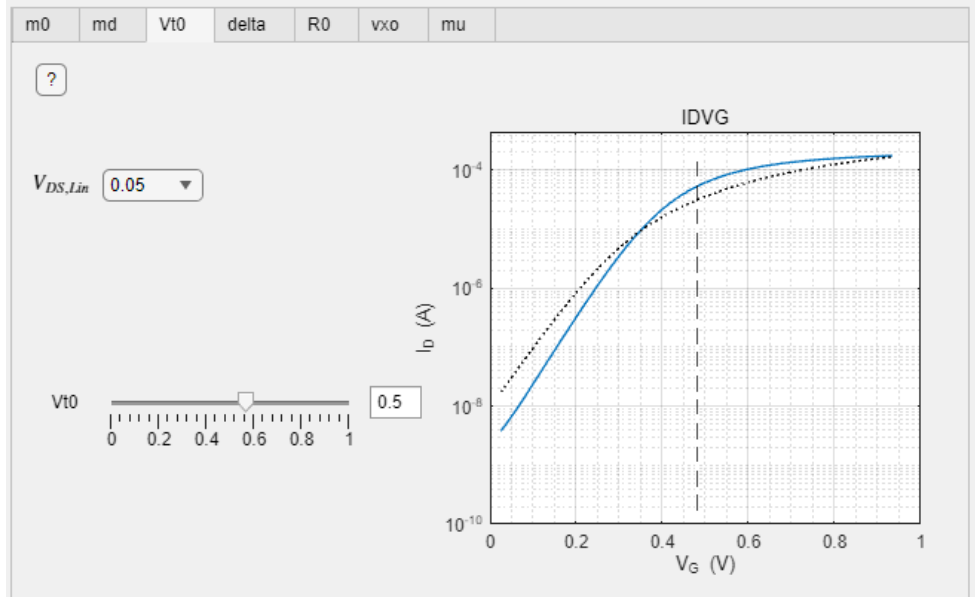
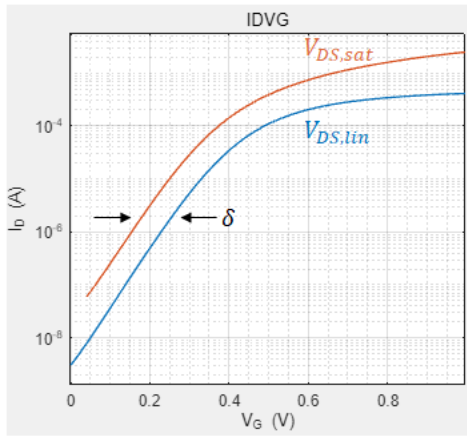


Figure 5.8:  $V_{th}$  extraction GUI for a single temperature.

a)



b)

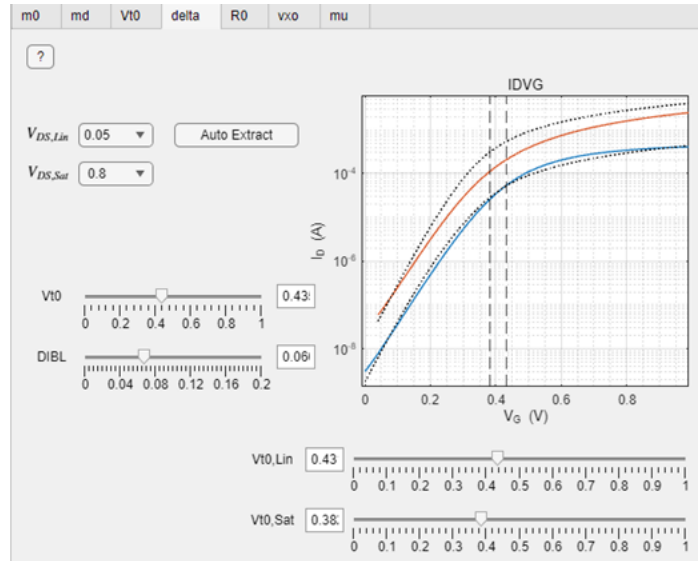


Figure 5.9: (a) Graphical illustration the basis of  $\delta$  auto extraction routine. (b)  $V_{T0}$  and  $\delta$  extraction GUI for a single temperature.

Moreover, a portion of  $V_{DS}'$  contributes to channel inversion, reducing  $V_{th}$  in a phenomenon known as DIBL. In a similar manner to initial  $m$ , the extracted  $V_{th}$  so far is not MVS's  $V_{T0}$  parameter, but with an added small contribution from  $V_{DS,lin}$  (eq: 4.17c). Therefore,  $V_{th}$  is modeled by two factors:  $V_{T0}$  and  $\delta$  (eq. 4.17d), requiring two transfer curves (typically in the linear and saturation regions), where  $\delta$  models the DIBL effect as (Fig. 5.9a):

$$\delta = \frac{\Delta V_{th}}{\Delta V_D} \quad (5.3)$$

Similar to step 1, the extraction process can be performed on either of two parameter pairs governed by the linear system:

$$\begin{bmatrix} V_{th,lin} \\ V_{th,sat} \end{bmatrix} = \begin{bmatrix} 1 & -V_{DS,lin} \\ 1 & -V_{DS,sat} \end{bmatrix} \cdot \begin{bmatrix} VT0 \\ \delta \end{bmatrix} \quad (5.4)$$

The flow at this stage will be similar to step 2 to that of extracting  $m_d$ , where the visual aids are based on the  $V_{th,sat}$  and  $V_{th,lin}$  pair, and their relation to the  $V_{T0}$  and  $\delta$  pair is automatically handled by the tool. The automation routine extracts  $\delta$  based on eq. 5.3.

### Step 3: Internal Resistance (Related Parameters: $R_0$ )

Unlike the previous steps, this extraction is most convenient from the output characteristics perspective (Fig. 5.10). In the linear region, the total resistance of the device can be extracted from the  $I_D$ - $V_D$  relationship:

$$R_{device} = \frac{\Delta V_{DS}}{\Delta I_D} = R_S + R_{ch} + R_D \quad (5.5a)$$

$$R_S = R_D = R_0/W \quad (5.5b)$$

where  $R_{ch}$  is the channel resistance,  $R_S$  and  $R_D$  are the source and drain resistances, respectively. For a symmetric device  $R_S = R_D$ . Since the channel resistance is dependant on bias conditions, this step can only predict  $R_0$  range. The best way to determine this range is when  $R_{ch}$  is at its lowest, which happens at maximum  $V_{GS}$  and minimum  $V_{DS}$ , as depletion would be minimized and channel inversion would be maximized, respectively. The tool auto selects the highest  $V_{GS}$  bias available, and offers the ability to change it.

The visual aid assumes  $R_{ch} = 0$ , showing the  $I - V$  relationship of  $R_{device} = 2R_0W$ , corresponding to the chosen gate voltage. Similar to step 1, the visual aid originated from a (VD0, ID0) which the user can modify. On the other hand, the principle behind the automatic extraction routine is to estimate  $R_{0,max}$  at the most favorable bias conditions, while assuming  $R_{ch} = 0$ :

$$R_{0,max} = \frac{\Delta V_{DS}}{\Delta I_D} / 2 \times W \quad (5.6)$$

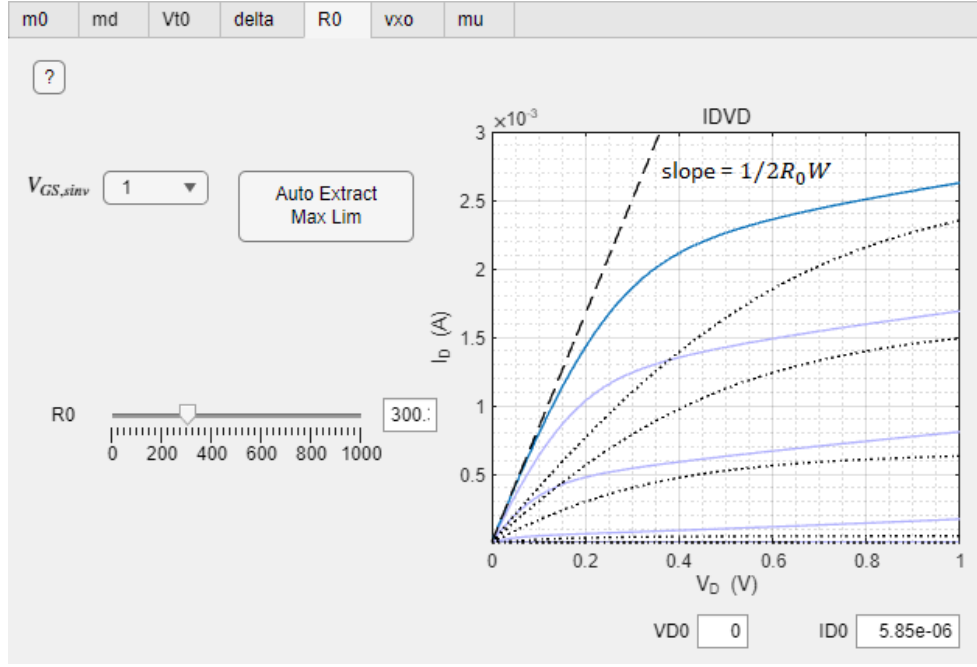


Figure 5.10:  $R_0$  extraction GUI for a single temperature based on a chosen  $I_D$ - $V_D$  curve in strong inversion. The chosen curve is highlighted, while the others are faded.

#### Step 4: Transport Parameters Extraction (Related Parameters: $\mu_{app}$ , $v_{xo}$ , $\beta$ )

The mobility used in the MVS is the low field apparent mobility, and hence, it is best fitted in the linear region where the  $I_D$ - $V_{GS}$  curve is fitted. From the  $I_D$ - $V_{GS}$  perspective, the effect is manifested in at low  $V_{DS}$  values, in conjunction with ring-oscillator (RO), affecting the point at which  $V_{DSAT}$  occurs:

$$V_{DSAT} = \frac{\mu_{app} L_{eff}}{v_{xo}} \quad (5.7)$$

The velocity is similarly extracted but with more emphasis on the situation region, where the velocity had saturated at  $v_{xo}$ . Therefore, the saturation transfer curve is fitted, while the effect on  $I_D$ - $V_{GS}$  is manifested as raising and lowering saturation  $I_D$  lines. This effect also seems in conjunction of  $R_0$ 's. It is also important to recognize that the slope of these saturation lines is heavily influenced by DIBL. Therefore, it is recommended to not hesitate to re-adjust previously extracted parameters (like  $R_0$  and  $\delta$ ) to obtain a decent fit.

Moreover, in rare occasions,  $\beta$  may need to be also adjusted, as it plays a major role here in the transition between linear and saturation regions (eq. 4.17e. Its effect is manifested on the

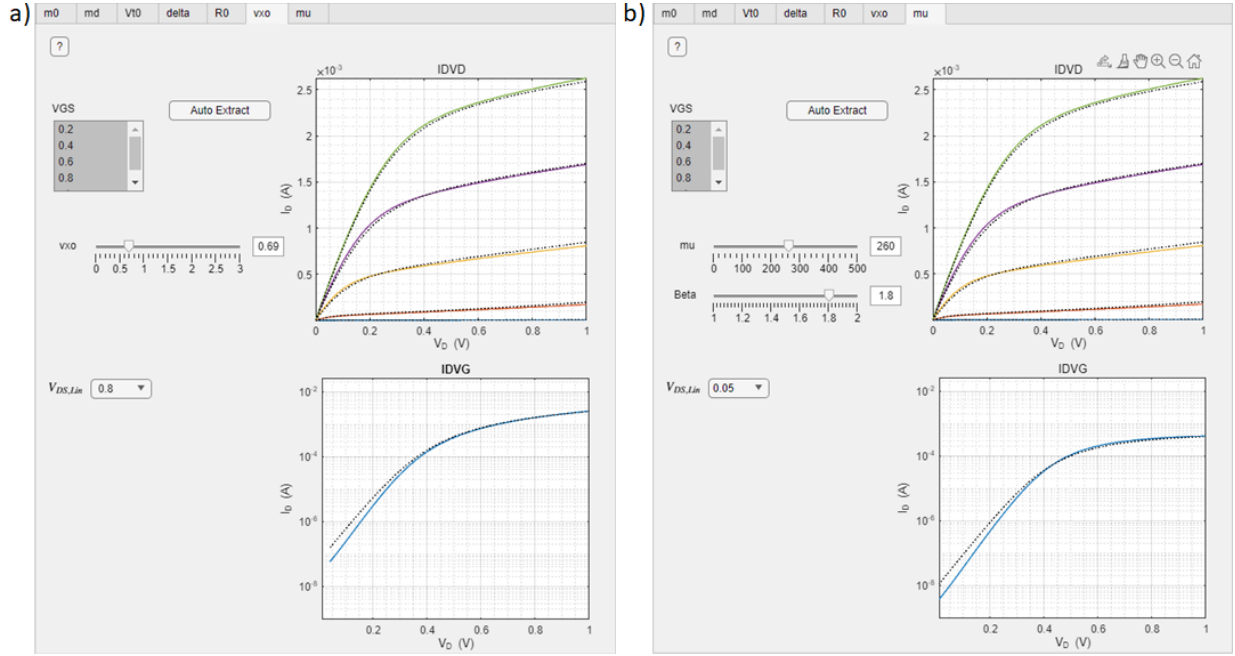


Figure 5.11: (a)  $\mu_{app}$  and (b)  $v_{xo}$  extraction GUI for a single temperature.

output characteristics by adjusting the roundness (or sharpness) of the kink at  $V_{DSAT}$ . Its range is limited to 1-2 with 1.8 being typical for NMOS and 1.6 for PMOS [79].

### Step 5: Model Fitting and Parameter Fine Tuning

After obtaining a rough overall fit, all parameters can then be fine-tuned for the final temperature point fit. This step is similar to base MVS parameter extraction, employing non-linear regression algorithms to simultaneously fit all parameters at once. However, these methods become more complex with more parameters, and can get stuck in local optimal values. In other words, the earlier steps were means for obtaining an initial guess with reasonable values, and help the auto-fitting routine to land on the global optimum solution.

In this tool, additional options are added to provide more control over the automatic regression. The sequence in which the  $I - V$  characteristics are alternated can be set by the user, along with the number of iterations desired. Moreover, the tool allows for selective exclusion of parameters per  $I - V$  iteration. For example, the parameters  $m_0$ ,  $V_{T0}$  are more pronounced in  $I_D - V_{GS}$  than  $I_D - V_{DS}$ , and hence, fitting  $I_D - V_{DS}$  can be except from adjusting them. The user may also differentiate this selection between the initial iteration and subsequent iterations. For instance,

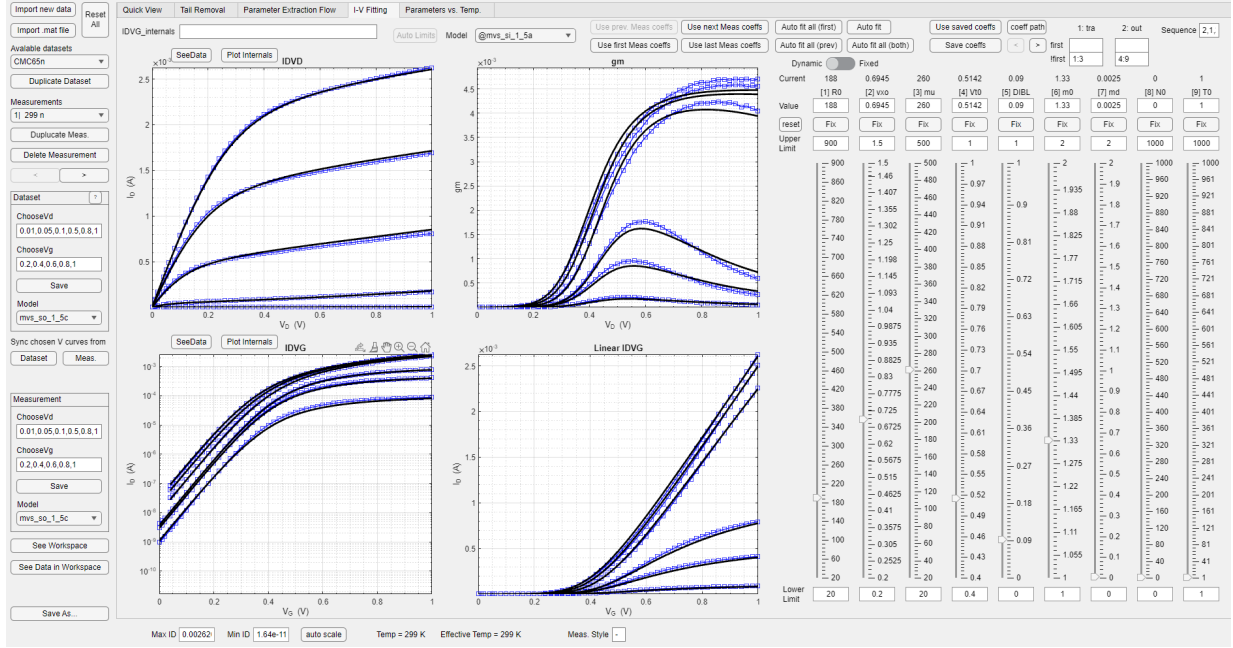


Figure 5.12: Model fitting GUI where all parameters can be simultaneously fine-tuned for one temperature point.

the settings shown in Fig. 5.12 are set for a sequence of  $I_D-V_{DS}$  first, then  $I_D-V_{GS}$ , then back to  $I_D-V_{DS}$ , and so on. In the first iteration, all parameters are adjusted according to  $I_D-V_{DS}$ , but  $R_0$ ,  $\mu$ , and  $v_{xo}$  are not adjusted in the second iteration with  $I_D-V_{GS}$ . Furthermore, subsequent  $I_D-V_{DS}$  iterations are also barred from adjusting any values other than  $R_0$ ,  $\mu$ ,  $v_{xo}$ , and  $V_{TO}$ .

As fitting for different temperatures progresses, the option to skip the parameter extraction flow and base the initial guess on adjacent temperature points is also available. Finally, in case the automatic regression does not produce acceptable results, manual fine-tuning is also supported via sliders and direct value injection.

### 5.2.3 Temperature Dependence Extraction

For  $T$ -dependent parameters, another suite of extraction tools is developed. Some of these tools are independent of the steps in 5.2.2, while others are reliant on the parameter sets of at least 2-3 temperature points. It is also possible to perform  $T$ -dependence extraction after all available  $T$ -points have been individually extracted, as a refining step to the final parameter set. Since more than one good fit can exist for each temperature point (in addition to factoring in the the

nature of experimental results and anomalous points), the option to exclude these temperature points, to focus the extraction process, is available for each of these tools.

Moreover, reload (or refresh) buttons are provided for each step to revert back to the last saved parameters, as well as a clear buttons to restart the extraction from scratch. Automatic extraction routines use the parameter values shown on the GUI as an initial guess, and will not necessarily adhere to them. To manually perform the extraction, it is possible to adjust these parameters individually. The tool will automatically update the graph and save the last entered parameters.

### Step 6: Subthreshold Swing Temperature Dependence

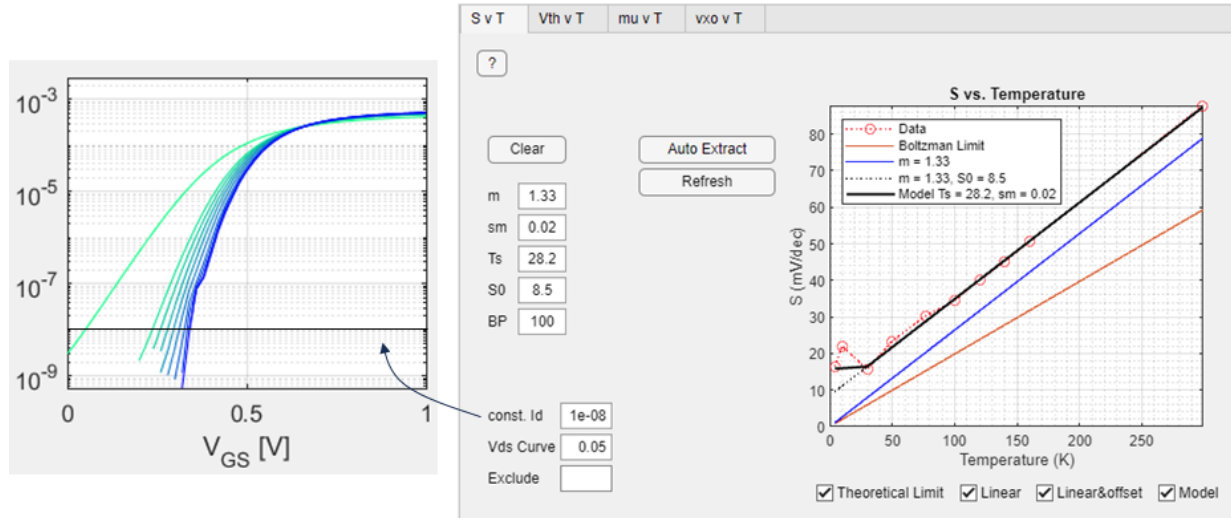


Figure 5.13:  $S$  vs.  $T$  extraction GUI for  $sm$ ,  $S_0$ , and  $T_S$ . The tool offers the option to change the  $V_{DS}$  bias of the transfer curve, as well as the constant current at which  $S$  is extracted.

This step focuses on extracting temperature characteristics of the subthreshold swing related to its offset and the effective temperature (eq. 4.21). The data preparation step is essential here since  $S$  is automatically extracted from raw measurement data-points. Failure to prepare the data, especially in the linear region, may result in erroneous  $S$  values or even crashing the tool.

The  $S$  information are typically extracted in the linear region when  $V_{DS} = 0.05V$ , however, it is possible to extract them from other  $V_{DS}$  biases. The extraction is performed at a constant current level for all temperatures, which should be carefully chosen, especially for bulk devices

with low subthreshold slopes at RT. Consider the linear transfer curves plotted for every measured temperature in Fig. 5.13b. While selecting a low current level is desirable to minimize the variable  $m$  effect, it could be too low for higher temperature measurements, leading to extrapolation and potential errors. By default the lowest current point at RT is chosen, but it should be carefully checked to minimize  $V_{GS}$  effect on  $m$ , especially at cryoT. To ensure accurate extraction, intermediate points are interpolated when the exact point is not present for any temperature curve.

The subthreshold swing factor can be extracted directly from measurement results using eq. 5.1.  $m$  here can be seen as the  $T$ -independent portion,  $m_{\mathcal{T}} = m_0 + V'_{DS}m_d$ , which was used in step 1. The extraction tool provides several useful curves to aid in analyzing the  $S(T)$  profile: the Boltzmann limit (eq. 5.8a where  $m = 1$ ), the classical profile (eq. 5.8b), the offset effect (eq. 5.8c), and the complete model at a constant  $V_{DS}$  bias (eq. 5.8d):

$$S_{limit}(T) = \phi_T \ln(10) \quad (5.8a)$$

$$S_{classic}(T) = m_{\mathcal{T}} \phi_T \ln(10) \quad (5.8b)$$

$$S_{w/offset}(T) = m_{\mathcal{T}} \phi_T \ln(10) + S_0 \quad (5.8c)$$

$$S(T)_{model} = m_{\mathcal{T}} \frac{\phi_T}{\phi_{T_e}} \ln(10) + S_0 \quad (5.8d)$$

Moreover, an extra parameter can be seen in Fig. 5.13 called the break-point (BP). BP is not a cryoMVS model parameter, but instead, an internal variable to the automation routine. The reason this variable is exposed in the GUI is because it may require user intervention in some cases. The extraction tool sees the  $S(T)$  profile as a piece-wise function and must detect the break point between the downward and horizontal trends. While automatic detection is implemented, the profile for some DUTs may not be well behaved, and user intervention may be needed. In such a case, BP must be set to a temperature point close to where the trends meet. It can be ignored otherwise, even if its value does not seem to match the  $S(T)$  profile, as long as  $T_S$  is reasonable. Finally, the  $V_{DS}$ -dependence parameters ( $N_o$  and  $T_o$ ) are simply the average of extracted values in step 5 across the available temperature range.

## Step 7: Threshold Voltage Temperature Dependence

The  $V_{th}(T)$  extraction process is among the simplest processes, where most of the complicated math is handled by the tool. The data-points shown in Fig. 5.14 are MVS's  $V_{T0}$  parameters fitted at multiple temperatures. For best results, it is recommended to first perform the steps in section 5.2.2 for the highest and lowest temperature points, as well as for  $T = \sim 50$  K.

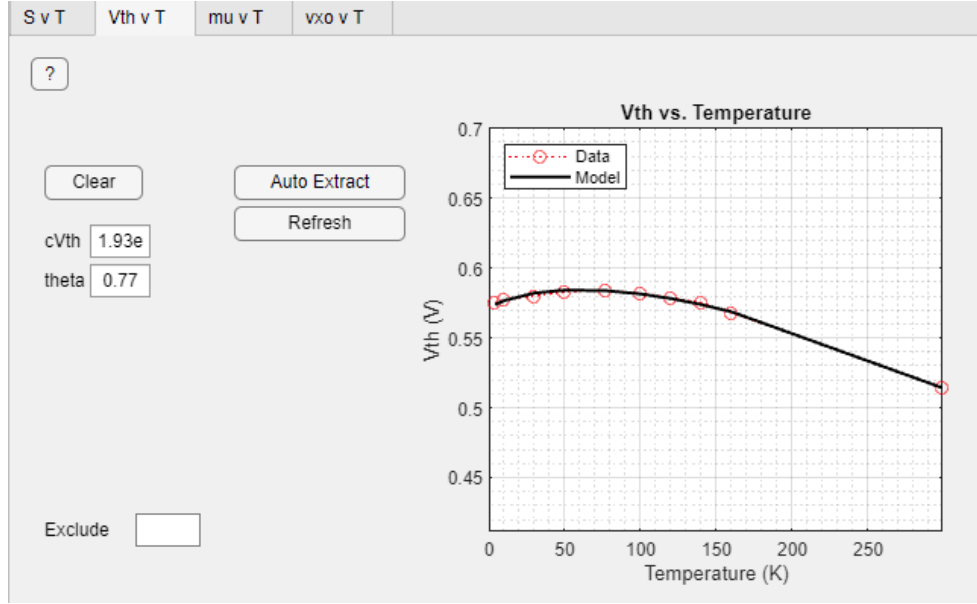


Figure 5.14:  $V_{th}$  vs.  $T$  extraction GUI for  $c_{V_{th}}$  and  $\theta$ .

### Step 8: Transport Parameters Temperature Dependence

The extraction process for both  $\mu_{app}$  and  $v_{inj}$  are identical because they use the same mathematical form to model temperature dependence (eqs. 4.23 and 4.24). Similar to step 7, the data-points shown in Figs. 5.15 and 5.16 are MVS's  $\mu$  and  $v_{xo}$  parameters, respectively, fitted at multiple temperatures. In a similar manner to step 6, the automation routine identifies the break point between the upward and downward trends, and in this case, the parameters  $T_{mu}$  and  $T_{vxo}$  are sufficient. If the automation routine struggles to extract the parameters, setting this break-point manually can help. For best results, it is recommended to first perform the steps in section 5.2.2 for the highest and lowest temperature points, as well as for a  $T$  point where the mobility and velocity trends are expected to change (this is highly technology dependent)

During automatic extraction, the parameter values for the lowest, highest, and mid temperatures are used to predict  $a_1$ ,  $a_2$ ,  $p_1$  and  $p_2$  as follows:

$$p_1 = \frac{\log_{10}(\mu_{min}/\mu_{mid})}{\log_{10}(T_{min}/T_{mid})} \quad (5.9a)$$

$$p_2 = \frac{\log_{10}(\mu_{mid}/\mu_{max})}{\log_{10}(T_{mid}/T_{max})} \quad (5.9b)$$

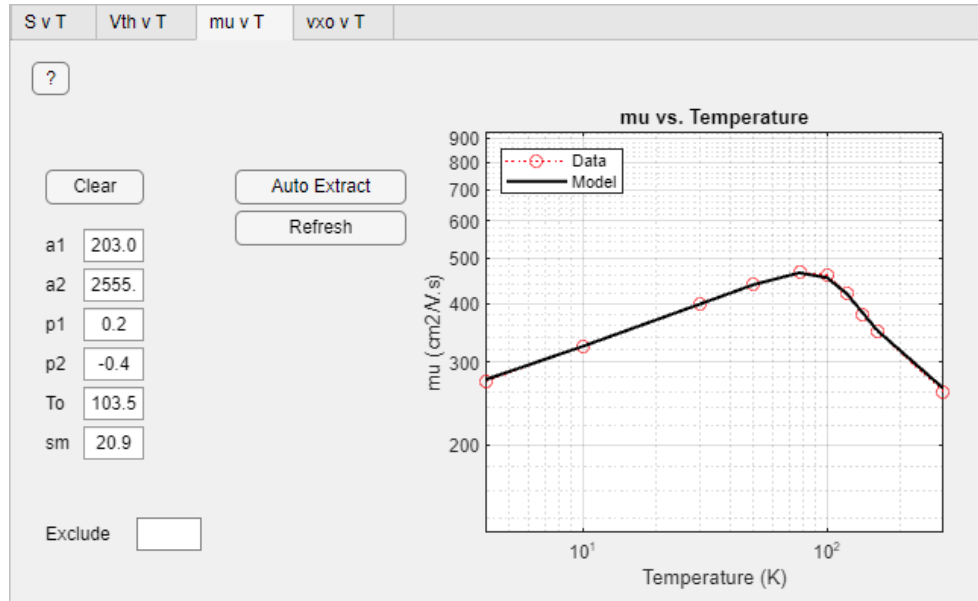


Figure 5.15:  $\mu_{app}$  vs.  $T$  extraction GUI for  $a_1$ ,  $a_2$ ,  $p_1$ ,  $p_2$ ,  $T_{mu}$ , and  $sm_{mu}$ .

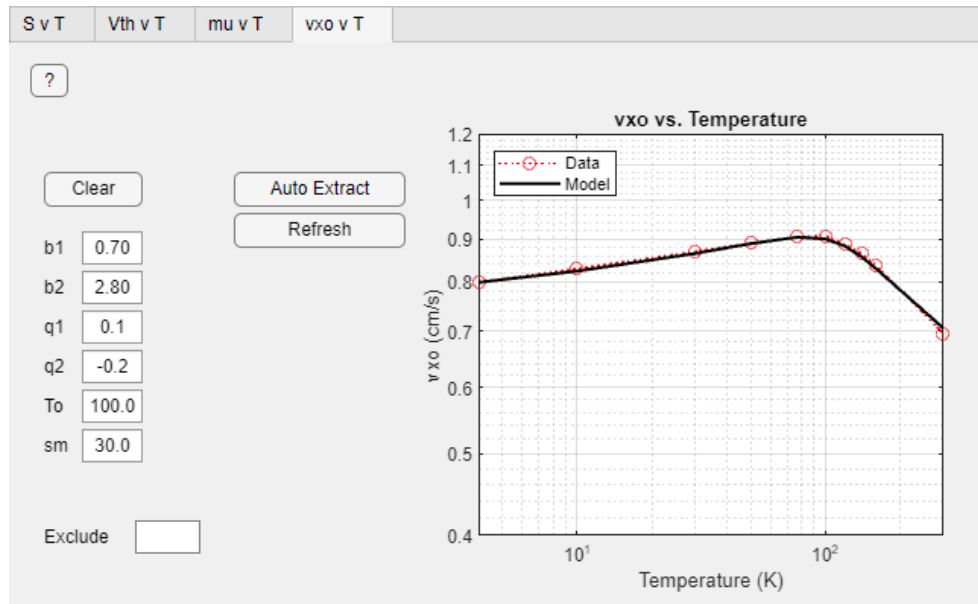


Figure 5.16:  $v_{inj}$  vs.  $T$  extraction GUI for  $b_1$ ,  $b_2$ ,  $q_1$ ,  $q_2$ ,  $T_{vxo}$ , and  $sm_{vxo}$ .

$$a_2 = \frac{\mu_{mid}}{T_{mid}^{p_2}} \quad (5.9d)$$

### 5.3 Circuit Simulation

Fig. 5.18 shows the CMC65n transistor output and transfer characteristics as interpreted by the circuit simulator through the cryoMVS model. With these models, transient circuit simulations at two temperature points (RT = 300 K and cryoT = 4 K) are conducted using Cadence EDA tools and Spectre circuit simulator as shown in Fig. 5.19. The results show successful operation at 200 MHz (a typical frequency for such applications) with potential for much faster operation for both temperatures. This showcases the ability for mature technologies to meet the demands

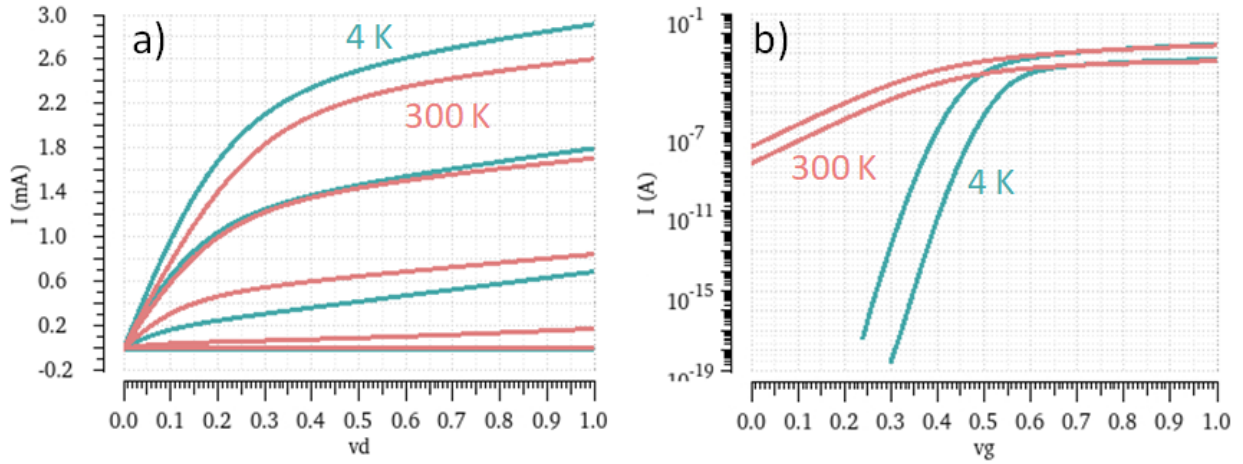


Figure 5.18: CMC65n  $I$ - $V$  characteristics as seen by Cadence. (a)  $I_D$ - $V_D$  and (b)  $I_D$ - $V_G$ .

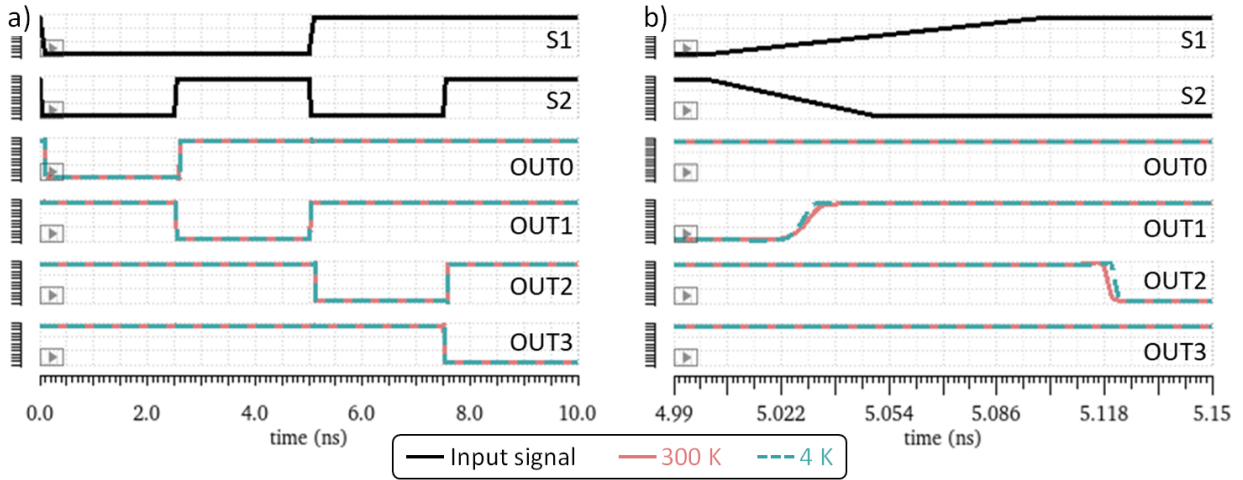


Figure 5.19: Transient circuit simulation of circuit Fig. 5.17 at 300 K and 4 K. D and EN are set to 1 V (or logical high). (a) Zoomed-out view of entire simulation time, (b) Zoomed-in view showing some signal transistors in detail.

of quantum computing applications. The model was not subjected to frequency stress-testing because it focuses on static parameters, and further development is required to improve dynamic operation accuracy.

A notable improvement in stand-by leakage is predicted by the simulation ( $\mu W$  to  $nW$ ) owing to the significant reduced leakage current at cryoT. Moreover, a slightly larger propagation delay

(the time needed for an input signal to propagate to the output of the logic gate) is observed at cryoT. This can be a result of the unequal increase in  $V_{th}$  between CMC65n and CMC65p ( $\sim .1$  V vs.  $\sim .2$  V) with temperature decrease, among other unequal parameter variation with  $T$ . Moreover, the circuit is originally designed for RT operation, and these unaccounted for parameter changes can affect its performance negatively. However, the transient delay (the rise or fall time of a signal) seem to improve with temperature decrease (2.4 ps vs. 1.8 ps). This is realistic because, to the first order, the overall increase in drive current at low temperatures would speed up capacitor charge and discharge, speeding up signal transition.

In summary, the successful run of the circuit simulation validates the model's practical functionality in real-world circuit design flows, and signifies a step forward towards realizing well-designed cryo-electronics.

# Chapter 6

## Conclusions and Future Research

### 6.1 Conclusions

In this thesis, the cryogenic operation of MOSFET is investigated to enable cryo-electronics design and development needed by a wide range of applications, like quantum computing. The investigation included experimental measurements at room temperature and cryogenic temperatures, revisiting the traditional physical model and expanding it to include operation at such extreme temperatures, the practical requirements and the development of a cryogenic compact model capable of aiding circuit designers and running in circuit simulation environments, and the implementation and packaging of the compact model in preparation for use by circuit designers.

In more detail, chapter 3 focused on cryogenic measurements of multiple samples from three distinct process nodes, ensuring comprehensive coverage of bulk CMOS technologies. Additional sample chip with known layout, ESD protection, easy access to internal structures, and an on-chip sensor was designed and taped-out utilizing industry standard EDA tools and PDK. Then, the chapter explored various options for conducting cryogenic measurements, meticulously addressing all potential interference issues. Two different cryostat options were investigated for initial and standard testing along with two different instruments setups that account for modularity, remote-access, and availability of enough SMUs. The cryogenic effects on the dipper probe were also investigated and compensated for. Moreover, sample preparation considerations were developed including a unique wire-bonding technique for stable yield. This extensive effort resulted in successful measurements of all samples with distinguishable properties that was the basis for the other aspects of this thesis.

In chapter 4 the fundamental physics for semiconductors and devices is revisited for cryogenic operation with particular focus on common assumptions and approximations. Specific

attention is paid to the  $E_F$  level, as numerical and analytical methods were employed to pinpoint its exact location within the band gap at various temperature and doping conditions. The semiconductor degeneracy and dopant ionization degree were brought to question and investigated at cryoT. It was found that MB approximation is valid in cold temperatures with great accuracy, but incomplete ionization is required for correct analysis. The chapter then developed temperature-dependence models for  $V_{th}$ ,  $S$ ,  $\mu_{app}$ , and  $v_{inj}$ . The latter two models were empirical in nature due to the sheer complexity of the scattering mechanisms involved, to reduce the number of parameters, and eliminate the requirement of proprietary knowledge and trade-secrets related to the technology process. The chapter also introduces the original MVS compact model and discusses its expansion to include cryoT. Each of the physical models are subsequently converted and optimized for practical circuit simulation. Finally, to showcase the model's capability, it was used to fit the  $I$ - $V$  characteristics measured in chapter 3.

In chapter 5, the compact model is prepared for employment in simulation environments used in the industry. It was implemented in Verilog-A where specific requirements and guidelines are reviewed, as well as testing. The compact model is also packaged with a suit of data preparation, parameter extraction, and model fitting tools. After describing each of the tools, the model is validated to work well with industry standard EDA tools and circuit simulators.

## 6.2 Contributions

The contributions made in this research can be summarized into five different items:

1. Exploring cryoT effect on device performance, including already reported phenomena and new discoveries. This work produced one of the first publications to report and model the  $V_{GS}$ -dependence phenomenon in  $S$ .
2. Developing unique models for the temperature dependence of key CMOS performance parameters (e.g.  $V_{th}$ ,  $S$ ,  $\mu$ ,  $v$ ) and incorporating them all into one unified compact model.
3. The only study that expands the MIT Virtual-Source (MVS) model to include cryoT. MVS is a flexible model that excels in technology benchmarking, evaluation, and facilitating first-pass circuit design simulations.
4. Providing parameter extraction and model fitting tools, a rarity for cryoT research.
5. Including circuit simulation validation of a relevant circuit design (Demux)

## 6.3 Future Research

Next, some future research venues are highlighted. With regards to CMOS technology, there is a sizable interest in SOI-MOSFET technologies in literature owing the greater control provided by the second gate and increased density, despite the higher costs. Even the more advanced FDSOI-MOSFET process nodes are entering maturity and on the verge of joining bulk MOSFET as economically viable options with high yields. Indeed, the additions made to create the cryoMVS model should be transferable to the FDSOI version of MVS with minimum effort. This transfer will still require validation against experimental results of FDSOI devices.

Furthermore, in the development of the cryoMVS model, the primary focus was placed on the device's DC (or static) characteristics, overlooking crucial analog performance parameters. These include transconductance ( $gm$ ), parasitic capacitance, and terminal charge. These time-dependent (or dynamic) parameters are essential for accurately predicting the circuit's behavior in time-dependent simulations such as the one presented in subsection 5.3. Indeed, the presented simulation relied on the dynamic RT model included in base MVS to simulate the dynamic behaviour, while the cryoT static model handled the static performance parameters. Thus, incorporating models that account for terminal-associated charges and capacitances becomes imperative to bridge this gap and enhance temporal accuracy in simulations. Given that MVS is charge-based model, it is recommended to conduct studies akin to [80] in cryoT conditions.

To improve the static model's accuracy, it is not recommended to spend the effort in lower temperature measurements approaching few mK. Instead, efforts must be geared towards cryogenic probing stations with tested thermal isolation and specified cryogenic properties. Cryogenic measurements require long periods of time to safely transition between a wide range of temperatures ( $\sim 300$  K) without damaging the DUTs's mechanical integrity. This amount of time must be taken into consideration while planning for such measurements, and redundancy can help greatly improve measurement throughput. Moreover, investing in improving the modularity of high-end semiconductor characterizers (like the B1500) outside of proprietary software would highly improve measurement convenience, time, and accuracy.

Regarding the physical model, analytical models for transport parameters is potentially possible for a smaller range of device technologies. Specific focus on interface traps profiling at cryoT is another possible research avenue. Moreover, there was a notable lack of experimental research on the values of metal work-function, electron affinity, vacuum and ionization energy levels in cryoT, which resulted in assuming them all constant in the analysis. Furthermore, important aspects of compact models that does not receive enough attention in cryogenic research are: stress testing, reliability, and variation. These aspects are crucial in today's circuit design industry, and a serious contributor to overall yield. They must be also investigated in cryoT and their temperature-dependence must be determined.

Finally, the CM's supporting tools can be further improved. This work focused on the GUI aspect to improve usability and adoption. However, the visual components of a software are difficult to maintain and produce unnecessary bugs and workload. Moreover, experienced circuit designers have long forgone the visual appeal of a software in favor for more robust, dynamic, and effective command-line and scripting tools. In the event that a GUI is needed, the common practice in the industry is to develop proprietary software that is optimized specifically to the in-house design flow. Therefore, instead of developing visual interface, robust modular routines that can be invoked by external software is a more effective method and recommended for any future upgrades on the cryoMVS model. Moreover, MATLAB is a powerful mathematical platform for CM development, but its GUI capabilities are found to be lacking. From slow response to the inability to support external calls by other programming languages (common in a professional circuit design environment), it is now recommended to use other platforms for such tools like Python or Perl. They are widely used in the industry and receive respectable support from the software community. An implementation in such languages would be easily implemented in an industry design flow, reducing the barrier-to-entry and improving the overall adoption of the compact model.

# References

- [1] *General Properties of Silicon*. en-US. URL: <https://www.pveducation.org/pvcdrom/materials/general-properties-of-silicon>.
- [2] Richard P Feynman. *Quantum mechanical computers*. en.
- [3] Ashley Montanaro. “Quantum algorithms: an overview”. en. In: *npj Quantum Information* 2.1 (Jan. 2016), p. 15023. ISSN: 2056-6387. DOI: [10.1038/npjqi.2015.23](https://doi.org/10.1038/npjqi.2015.23). URL: <http://arxiv.org/abs/1511.04206>.
- [4] Lieven Vandersypen and Antoni van Leeuwenhoek. “1.4 Quantum computing - the next challenge in circuit and system design”. en. In: *2017 IEEE International Solid-State Circuits Conference (ISSCC)*. San Francisco, CA, USA: IEEE, Feb. 2017, pp. 24–29. ISBN: 978-1-5090-3758-2. DOI: [10.1109/ISSCC.2017.7870244](https://doi.org/10.1109/ISSCC.2017.7870244). URL: <http://ieeexplore.ieee.org/document/7870244/>.
- [5] *On the Path to Cryogenic Control of Quantum Processors*. en. Feb. 2019. URL: <https://ai.googleblog.com/2019/02/on-path-to-cryogenic-control-of-quantum.html>.
- [6] Frank Arute et al. “Quantum supremacy using a programmable superconducting processor”. en. In: *Nature* 574.7779 (Oct. 2019), pp. 505–510. ISSN: 0028-0836, 1476-4687. DOI: [10.1038/s41586-019-1666-5](https://doi.org/10.1038/s41586-019-1666-5). URL: <http://www.nature.com/articles/s41586-019-1666-5>.
- [7] *On “Quantum Supremacy”*. en-US. Oct. 2019. URL: <https://web.archive.org/web/20230215010715/https://www.ibm.com/blogs/research/2019/10/on-quantum-supremacy/>.
- [8] N H Balshaw. *Practical Cryogenics*. en.
- [9] Bishnu Patra et al. “Cryo-CMOS Circuits and Systems for Quantum Computing Applications”. en. In: *IEEE Journal of Solid-State Circuits* 53.1 (Jan. 2018), pp. 309–321. ISSN: 0018-9200, 1558-173X. DOI: [10.1109/JSSC.2017.2737549](https://doi.org/10.1109/JSSC.2017.2737549). URL: <https://ieeexplore.ieee.org/document/8036394/>.

- [10] S. Ramesh Ekanayake et al. “Characterization of SOS-CMOS FETs at Low Temperatures for the Design of Integrated Circuits for Quantum Bit Control and Readout”. en. In: *IEEE Transactions on Electron Devices* 57.2 (Feb. 2010), pp. 539–547. ISSN: 0018-9383, 1557-9646. DOI: [10.1109/TED.2009.2037381](https://doi.org/10.1109/TED.2009.2037381). URL: <http://ieeexplore.ieee.org/document/5373940/>.
- [11] J.M. Hornibrook et al. “Cryogenic Control Architecture for Large-Scale Quantum Computing”. en. In: *Physical Review Applied* 3.2 (Feb. 2015), p. 024010. ISSN: 2331-7019. DOI: [10.1103/PhysRevApplied.3.024010](https://doi.org/10.1103/PhysRevApplied.3.024010). URL: <https://link.aps.org/doi/10.1103/PhysRevApplied.3.024010>.
- [12] E. Charbon et al. “Cryo-CMOS for quantum computing”. en. In: *2016 IEEE International Electron Devices Meeting (IEDM)*. San Francisco, CA, USA: IEEE, Dec. 2016, pp. 13.5.1–13.5.4. ISBN: 978-1-5090-3902-9. DOI: [10.1109/IEDM.2016.7838410](https://doi.org/10.1109/IEDM.2016.7838410). URL: <http://ieeexplore.ieee.org/document/7838410/>.
- [13] Eduardo B Ramirez et al. “Few-electrode design for silicon MOS quantum dots”. en. In: *Semiconductor Science and Technology* 35.1 (Jan. 2020), p. 015002. ISSN: 0268-1242, 1361-6641. DOI: [10.1088/1361-6641/ab516a](https://doi.org/10.1088/1361-6641/ab516a). URL: <https://iopscience.iop.org/article/10.1088/1361-6641/ab516a>.
- [14] T. F. Watson et al. “A programmable two-qubit quantum processor in silicon”. en. In: *Nature* 555.7698 (Mar. 2018), pp. 633–637. ISSN: 0028-0836, 1476-4687. DOI: [10.1038/nature25766](https://doi.org/10.1038/nature25766). URL: <http://www.nature.com/articles/nature25766>.
- [15] P. Galy et al. “Cryogenic Temperature Characterization of a 28-nm FD-SOI Dedicated Structure for Advanced CMOS and Quantum Technologies Co-Integration”. en. In: *IEEE Journal of the Electron Devices Society* 6 (2018), pp. 594–600. ISSN: 2168-6734. DOI: [10.1109/JEDS.2018.2828465](https://doi.org/10.1109/JEDS.2018.2828465). URL: <https://ieeexplore.ieee.org/document/8370029/>.
- [16] Alessandro Crippa et al. “Level Spectrum and Charge Relaxation in a Silicon Double Quantum Dot Probed by Dual-Gate Reflectometry”. en. In: *Nano Letters* 17.2 (Feb. 2017), pp. 1001–1006. ISSN: 1530-6984, 1530-6992. DOI: [10.1021/acs.nanolett.6b04354](https://doi.org/10.1021/acs.nanolett.6b04354). URL: <https://pubs.acs.org/doi/10.1021/acs.nanolett.6b04354>.
- [17] Andrea Corna et al. “Electrically driven electron spin resonance mediated by spin–valley–orbit coupling in a silicon quantum dot”. en. In: *npj Quantum Information* 4.1 (Feb. 2018), p. 6. ISSN: 2056-6387. DOI: [10.1038/s41534-018-0059-1](https://doi.org/10.1038/s41534-018-0059-1). URL: <https://www.nature.com/articles/s41534-018-0059-1>.

- [18] S. Rochette et al. “Quantum dots with split enhancement gate tunnel barrier control”. en. In: *Applied Physics Letters* 114.8 (Feb. 2019), p. 083101. ISSN: 0003-6951, 1077-3118. DOI: [10.1063/1.5091111](https://doi.org/10.1063/1.5091111). URL: <http://aip.scitation.org/doi/10.1063/1.5091111>.
- [19] Kevin Eng et al. “Isotopically enhanced triple-quantum-dot qubit”. en. In: *Science Advances* 1.4 (May 2015), e1500214. ISSN: 2375-2548. DOI: [10.1126/sciadv.1500214](https://doi.org/10.1126/sciadv.1500214). URL: <https://www.science.org/doi/10.1126/sciadv.1500214>.
- [20] Christian Enz, Maria-Anna Chalkiadaki, and Anurag Mangla. “Low-power analog/RF circuit design based on the inversion coefficient”. en. In: *ESSCIRC Conference 2015 - 41st European Solid-State Circuits Conference (ESSCIRC)*. Graz, Austria: IEEE, Sept. 2015, pp. 202–208. DOI: [10.1109/ESSCIRC.2015.7313863](https://doi.org/10.1109/ESSCIRC.2015.7313863). URL: <http://ieeexplore.ieee.org/document/7313863/>.
- [21] Rosario M. Incandela et al. “Characterization and Compact Modeling of Nanometer CMOS Transistors at Deep-Cryogenic Temperatures”. en. In: *IEEE Journal of the Electron Devices Society* 6 (2018), pp. 996–1006. ISSN: 2168-6734. DOI: [10.1109/JEDS.2018.2821763](https://doi.org/10.1109/JEDS.2018.2821763). URL: <https://ieeexplore.ieee.org/document/8329135/>.
- [22] Arnout Beckers et al. “Cryogenic characterization of 28 nm bulk CMOS technology for quantum computing”. en. In: *2017 47th European Solid-State Device Research Conference (ESSDERC)*. Leuven, Belgium: IEEE, Sept. 2017, pp. 62–65. ISBN: 978-1-5090-5978-2. DOI: [10.1109/ESSDERC.2017.8066592](https://doi.org/10.1109/ESSDERC.2017.8066592). URL: <http://ieeexplore.ieee.org/document/8066592/>.
- [23] H. Achour et al. “Performances under saturation operation of p-channel FinFETs on SOI substrates at cryogenic temperature”. en. In: *2014 International Semiconductor Conference (CAS)*. Sinaia, Romania: IEEE, Oct. 2014, pp. 181–184. DOI: [10.1109/SMICND.2014.6966429](https://doi.org/10.1109/SMICND.2014.6966429). URL: <http://ieeexplore.ieee.org/document/6966429/>.
- [24] Arnout Beckers, Farzan Jazaeri, and Christian Enz. “Cryogenic MOS Transistor Model”. en. In: *IEEE Transactions on Electron Devices* 65.9 (Sept. 2018), pp. 3617–3625. ISSN: 0018-9383, 1557-9646. DOI: [10.1109/TED.2018.2854701](https://doi.org/10.1109/TED.2018.2854701). URL: <https://ieeexplore.ieee.org/document/8424046/>.
- [25] I. M. Hafez, G. Ghibaudo, and F. Balestra. “Assessment of interface state density in silicon metal-oxide-semiconductor transistors at room, liquid-nitrogen, and liquid-helium temperatures”. en. In: *Journal of Applied Physics* 67.4 (Feb. 1990), pp. 1950–1952. ISSN: 0021-8979, 1089-7550. DOI: [10.1063/1.345572](https://doi.org/10.1063/1.345572). URL: <http://aip.scitation.org/doi/10.1063/1.345572>.

- [26] Wriddhi Chakraborty et al. “An Empirically Validated Virtual Source FET Model for Deeply Scaled Cool CMOS”. en. In: *2019 IEEE International Electron Devices Meeting (IEDM)*. San Francisco, CA, USA: IEEE, Dec. 2019, pp. 39.4.1–39.4.4. ISBN: 978-1-72814-032-2. DOI: [10.1109/IEDM19573.2019.8993666](https://doi.org/10.1109/IEDM19573.2019.8993666). URL: <https://ieeexplore.ieee.org/document/8993666/>.
- [27] J. M. Boter et al. “A sparse spin qubit array with integrated control electronics”. en. In: *2019 IEEE International Electron Devices Meeting (IEDM)*. San Francisco, CA, USA: IEEE, Dec. 2019, pp. 31.4.1–31.4.4. ISBN: 978-1-72814-032-2. DOI: [10.1109/IEDM19573.2019.8993570](https://doi.org/10.1109/IEDM19573.2019.8993570). URL: <https://ieeexplore.ieee.org/document/8993570/>.
- [28] Hazem Elgabra et al. “Virtual Source based I-V Model for Cryogenic CMOS Devices”. en. In: *2019 International Symposium on VLSI Technology, Systems and Application (VLSI-TSA)*. Hsinchu, Taiwan: IEEE, Apr. 2019, pp. 1–2. ISBN: 978-1-72810-942-8. DOI: [10.1109/VLSI-TSA.2019.8804661](https://doi.org/10.1109/VLSI-TSA.2019.8804661). URL: <https://ieeexplore.ieee.org/document/8804661/>.
- [29] K. Eng et al. “Steps Towards Fabricating Cryogenic CMOS Compatible Single Electron Devices”. en. In: *2008 8th IEEE Conference on Nanotechnology*. Arlington, Texas, USA: IEEE, Aug. 2008, pp. 496–499. ISBN: 978-1-4244-2103-9. DOI: [10.1109/NANO.2008.149](https://doi.org/10.1109/NANO.2008.149). URL: <http://ieeexplore.ieee.org/document/4617131/>.
- [30] S. O’uchi et al. “Cryogenic operation of double-gate FinFET and demonstration of analog circuit at 4.2K”. en. In: *2012 IEEE International SOI Conference (SOI)*. Napa, CA, USA: IEEE, Oct. 2012, pp. 1–2. DOI: [10.1109/SOI.2012.6404376](https://doi.org/10.1109/SOI.2012.6404376). URL: <http://ieeexplore.ieee.org/document/6404376/>.
- [31] Laurent Varizat et al. “A low temperature 0.35um CMOS technology BSIM3.3 model for space instrumentation: Application to a voltage reference design”. en. In: *2017 IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace)*. Padua, Italy: IEEE, June 2017, pp. 74–78. ISBN: 978-1-5090-4234-0. DOI: [10.1109/MetroAeroSpace.2017.7999541](https://doi.org/10.1109/MetroAeroSpace.2017.7999541). URL: <http://ieeexplore.ieee.org/document/7999541/>.
- [32] M. Shin et al. “Low temperature characterization of 14nm FDSOI CMOS devices”. en. In: *2014 11th International Workshop on Low Temperature Electronics (WOLTE)*. Grenoble, France: IEEE, July 2014, pp. 29–32. ISBN: 978-1-4799-4842-0. DOI: [10.1109/WOLTE.2014.6881018](https://doi.org/10.1109/WOLTE.2014.6881018). URL: <http://ieeexplore.ieee.org/document/6881018/>.
- [33] Emil Arnold. “Disorder-induced carrier localization in silicon surface inversion layers”. en. In: *Applied Physics Letters* 25.12 (Dec. 1974), pp. 705–707. ISSN: 0003-6951, 1077-3118. DOI: [10.1063/1.1655369](https://doi.org/10.1063/1.1655369). URL: <http://aip.scitation.org/doi/10.1063/1.1655369>.

- [34] A. B. M. Hasan Talukder et al. “Temperature-Dependent Characteristics and Electrostatic Threshold Voltage Tuning of Accumulated Body MOSFETs”. In: *IEEE Transactions on Electron Devices* 69.8 (Aug. 2022), pp. 4138–4143. ISSN: 1557-9646. DOI: [10.1109/TED.2022.3184906](https://doi.org/10.1109/TED.2022.3184906).
- [35] B Cretu et al. “Assessment of DC and low-frequency noise performances of triple-gate FinFETs at cryogenic temperatures”. en. In: *Semiconductor Science and Technology* 31.12 (Dec. 2016), p. 124006. ISSN: 0268-1242, 1361-6641. DOI: [10.1088/0268-1242/31/12/124006](https://doi.org/10.1088/0268-1242/31/12/124006). URL: <https://iopscience.iop.org/article/10.1088/0268-1242/31/12/124006>.
- [36] Mark S. Lundstrom and Dimitri A. Antoniadis. “Compact Models and the Physics of Nanoscale FETs”. en. In: *IEEE Transactions on Electron Devices* 61.2 (Feb. 2014), pp. 225–233. ISSN: 0018-9383, 1557-9646. DOI: [10.1109/TED.2013.2283253](https://doi.org/10.1109/TED.2013.2283253). URL: <http://ieeexplore.ieee.org/document/6626559/>.
- [37] *Models – BSIM Group*. en-US. URL: <https://bsim.berkeley.edu/models/>.
- [38] *BSIM3 – BSIM Group*. en-US. URL: <http://bsim.berkeley.edu/models/bsim3/>.
- [39] Ybe Creten et al. “An 8-Bit Flash Analog-to-Digital Converter in Standard CMOS Technology Functional From 4.2 K to 300 K”. en. In: *IEEE Journal of Solid-State Circuits* 44.7 (July 2009), pp. 2019–2025. ISSN: 0018-9200. DOI: [10.1109/JSSC.2009.2021918](https://doi.org/10.1109/JSSC.2009.2021918). URL: <http://ieeexplore.ieee.org/document/5109798/>.
- [40] Burak Okcan et al. “A cryogenic analog to digital converter operating from 300 K down to 4.4 K”. en. In: *Review of Scientific Instruments* 81.2 (Feb. 2010), p. 024702. ISSN: 0034-6748, 1089-7623. DOI: [10.1063/1.3309825](https://doi.org/10.1063/1.3309825). URL: <http://aip.scitation.org/doi/10.1063/1.3309825>.
- [41] R. V. Langevelde and D. Klaassen. “Physical Background of MOS Model 11”. In: *Physical Background of MOS Model 11*. 2003. URL: <https://www.semanticscholar.org/paper/Physical-Background-of-MOS-Model-11-Langevelde-Klaassen/28db999cc5d3e313a57c33df769b904f330f0e86>.
- [42] G Gildenblat et al. *Introduction to PSP MOSFET Model (Invited)*. en.
- [43] *EKV MOSFET Model*. en-GB. URL: <https://www.epfl.ch/labs/iclab/ekv/>.
- [44] *Group: NEEDS: New Era Electronic Devices and Systems ~ MIT MVS Models*. URL: [https://nanohub.org/groups/needs/mvs\\_models](https://nanohub.org/groups/needs/mvs_models).

- [45] Germano S. Fonseca, Leonardo B. de Sá, and Antonio C. Mesquita. “Extraction of static parameters to extend the EKV model to cryogenic temperatures”. en. In: *Extraction of static parameters to extend the EKV model to cryogenic temperatures*. Ed. by Bjørn F. Andresen et al. Baltimore, Maryland, United States, May 2016, 98192B. DOI: [10.1117/12.2219734](https://doi.org/10.1117/12.2219734). URL: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2219734>.
- [46] Arnout Beckers, Farzan Jazaeri, and ChristianENZ. “Characterization and Modeling of 28-nm Bulk CMOS Technology Down to 4.2 K”. en. In: *IEEE Journal of the Electron Devices Society* 6 (2018), pp. 1007–1018. ISSN: 2168-6734. DOI: [10.1109/JEDS.2018.2817458](https://doi.org/10.1109/JEDS.2018.2817458). URL: <https://ieeexplore.ieee.org/document/8326483/>.
- [47] Arnout Beckers, Farzan Jazaeri, and ChristianENZ. “Theoretical Limit of Low Temperature Subthreshold Swing in Field-Effect Transistors”. en. In: *IEEE Electron Device Letters* 41.2 (Feb. 2020), pp. 276–279. ISSN: 0741-3106, 1558-0563. DOI: [10.1109/LED.2019.2963379](https://doi.org/10.1109/LED.2019.2963379). URL: <https://ieeexplore.ieee.org/document/8946710/>.
- [48] L. Deferm, E. Simoen, and C. Claeys. “The importance of the internal bulk-source potential on the low temperature kink in NMOSTs”. en. In: *IEEE Transactions on Electron Devices* 38.6 (June 1991), pp. 1459–1466. ISSN: 00189383. DOI: [10.1109/16.81639](https://doi.org/10.1109/16.81639). URL: <http://ieeexplore.ieee.org/document/81639/>.
- [49] R. M. Incandela et al. “Nanometer CMOS characterization and compact modeling at deep-cryogenic temperatures”. en. In: *2017 47th European Solid-State Device Research Conference (ESSDERC)*. Leuven, Belgium: IEEE, Sept. 2017, pp. 58–61. ISBN: 978-1-5090-5978-2. DOI: [10.1109/ESSDERC.2017.8066591](https://doi.org/10.1109/ESSDERC.2017.8066591). URL: <http://ieeexplore.ieee.org/document/8066591/>.
- [50] Chao Luo et al. “MOSFET characterization and modeling at cryogenic temperatures”. en. In: *Cryogenics* 98 (Mar. 2019), pp. 12–17. ISSN: 00112275. DOI: [10.1016/j.cryogenics.2018.12.009](https://doi.org/10.1016/j.cryogenics.2018.12.009). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0011227518301413>.
- [51] A. Akturk et al. “Compact and Distributed Modeling of Cryogenic Bulk MOSFET Operation”. en. In: *IEEE Transactions on Electron Devices* 57.6 (June 2010), pp. 1334–1342. ISSN: 0018-9383. DOI: [10.1109/TED.2010.2046458](https://doi.org/10.1109/TED.2010.2046458). URL: <http://ieeexplore.ieee.org/document/5456141/>.
- [52] Akin Akturk et al. “Device Modeling at Cryogenic Temperatures: Effects of Incomplete Ionization”. en. In: *IEEE Transactions on Electron Devices* 54.11 (Nov. 2007), pp. 2984–2990. ISSN: 0018-9383. DOI: [10.1109/TED.2007.906966](https://doi.org/10.1109/TED.2007.906966). URL: <http://ieeexplore.ieee.org/document/4367583/>.

- [53] A. Akturk et al. “Compact modeling of 0.35 $\mu$ m SOI CMOS technology node for 4K DC operation using Verilog-A”. en. In: *Microelectronic Engineering* 87.12 (Dec. 2010), pp. 2518–2524. ISSN: 01679317. DOI: [10.1016/j.mee.2010.06.005](https://doi.org/10.1016/j.mee.2010.06.005). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0167931710001917>.
- [54] Sujit K. Singh et al. “Accurate Modeling of Cryogenic Temperature Effects in 10-nm Bulk CMOS FinFETs Using the BSIM-CMG Model”. en. In: *IEEE Electron Device Letters* 43.5 (May 2022), pp. 689–692. ISSN: 0741-3106, 1558-0563. DOI: [10.1109/LED.2022.3158495](https://doi.org/10.1109/LED.2022.3158495). URL: <https://ieeexplore.ieee.org/document/9732438/>.
- [55] Girish Pahwa et al. “Compact Modeling of Temperature Effects in FDSOI and FinFET Devices Down to Cryogenic Temperatures”. en. In: *IEEE Transactions on Electron Devices* 68.9 (Sept. 2021), pp. 4223–4230. ISSN: 0018-9383, 1557-9646. DOI: [10.1109/TED.2021.3097971](https://doi.org/10.1109/TED.2021.3097971). URL: <https://ieeexplore.ieee.org/document/9497363/>.
- [56] P. Martin et al. “EKV3 compact modeling of MOS transistors from a 0.18 $\mu$ m CMOS technology for mixed analog–digital circuit design at low temperature”. en. In: *Cryogenics* 49.11 (Nov. 2009), pp. 595–598. ISSN: 00112275. DOI: [10.1016/j.cryogenics.2008.12.005](https://doi.org/10.1016/j.cryogenics.2008.12.005). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0011227508002129>.
- [57] R. Wang et al. “Threshold voltage variations with temperature in MOS transistors”. en. In: *IEEE Transactions on Electron Devices* 18.6 (June 1971), pp. 386–388. ISSN: 0018-9383. DOI: [10.1109/T-ED.1971.17207](https://doi.org/10.1109/T-ED.1971.17207). URL: <http://ieeexplore.ieee.org/document/1476529/>.
- [58] Nguyen Cong Dao et al. “An enhanced MOSFET threshold voltage model for the 6–300 K temperature range”. en. In: *Microelectronics Reliability* 69 (Feb. 2017), pp. 36–39. ISSN: 00262714. DOI: [10.1016/j.microrel.2016.12.007](https://doi.org/10.1016/j.microrel.2016.12.007). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0026271416304437>.
- [59] Arnout Beckers et al. “Physical Model of Low-Temperature to Cryogenic Threshold Voltage in MOSFETs”. en. In: *IEEE Journal of the Electron Devices Society* 8 (2020), pp. 780–788. ISSN: 2168-6734. DOI: [10.1109/JEDS.2020.2989629](https://doi.org/10.1109/JEDS.2020.2989629). URL: <https://ieeexplore.ieee.org/document/9076206/>.
- [60] S.K. Tewksbury. “Attojoule MOSFET logic devices using low voltage swings and low temperature”. en. In: *Solid-State Electronics* 28.3 (Mar. 1985), pp. 255–276. ISSN: 00381101. DOI: [10.1016/0038-1101\(85\)90006-1](https://doi.org/10.1016/0038-1101(85)90006-1). URL: <https://linkinghub.elsevier.com/retrieve/pii/0038110185900061>.

- [61] Arnout Beckers, Farzan Jazaeri, and Christian Enz. “Inflection Phenomenon in Cryogenic MOSFET Behavior”. en. In: *IEEE Transactions on Electron Devices* 67.3 (Mar. 2020), pp. 1357–1360. ISSN: 0018-9383, 1557-9646. DOI: [10.1109/TED.2020.2965475](https://doi.org/10.1109/TED.2020.2965475). URL: <https://ieeexplore.ieee.org/document/8974242/>.
- [62] H. Bohuslavskyi et al. “Cryogenic Subthreshold Swing Saturation in FD-SOI MOSFETs Described With Band Broadening”. en. In: *IEEE Electron Device Letters* 40.5 (May 2019), pp. 784–787. ISSN: 0741-3106, 1558-0563. DOI: [10.1109/LED.2019.2903111](https://doi.org/10.1109/LED.2019.2903111). URL: <https://ieeexplore.ieee.org/document/8660508/>.
- [63] Minju Shin et al. “In depth characterization of electron transport in 14 nm FD-SOI CMOS devices”. en. In: *Solid-State Electronics* 112 (Oct. 2015), pp. 13–18. ISSN: 00381101. DOI: [10.1016/j.sse.2015.02.012](https://doi.org/10.1016/j.sse.2015.02.012). URL: <https://linkinghub.elsevier.com/retrieve/pii/S003811011500057X>.
- [64] Antoine Cros et al. “Unexpected mobility degradation for very short devices : A new challenge for CMOS scaling”. en. In: *2006 International Electron Devices Meeting*. San Francisco, CA, USA: IEEE, 2006, pp. 1–4. ISBN: 978-1-4244-0438-4. DOI: [10.1109/IEDM.2006.346872](https://doi.org/10.1109/IEDM.2006.346872). URL: <http://ieeexplore.ieee.org/document/4154291/>.
- [65] J. R. Hoff et al. “Cryogenic Lifetime Studies of 130 nm and 65 nm nMOS Transistors for High-Energy Physics Experiments”. en. In: *IEEE Transactions on Nuclear Science* 62.3 (June 2015), pp. 1255–1261. ISSN: 0018-9499, 1558-1578. DOI: [10.1109/TNS.2015.2433793](https://doi.org/10.1109/TNS.2015.2433793). URL: <http://ieeexplore.ieee.org/document/7118268/>.
- [66] Arnout Beckers, Farzan Jazaeri, and Christian Enz. *Revised theoretical limit of subthreshold swing in field-effect transistors*. en. Dec. 2019. URL: <http://arxiv.org/abs/1811.09146>.
- [67] Rubaya Absar et al. “Integrated Control Addressing Circuits for a Surface Code Quantum Computer in Silicon”. In: *Proceedings of the 17th ACM International Symposium on Nanoscale Architectures*. NANOARCH ’22. New York, NY, USA: Association for Computing Machinery, May 2023, pp. 1–6. ISBN: 978-1-4503-9938-8. DOI: [10.1145/3565478.3572541](https://doi.org/10.1145/3565478.3572541). URL: <https://doi.org/10.1145/3565478.3572541>.
- [68] *Facts about Helium - Introduction — Quantum Design*. URL: <https://qd-europe.com/hu/en/product/facts-about-helium-introduction/>.
- [69] *TSMC 65 nm GP CMOS Process Technology – CMC Microsystems*. URL: <https://www.cmc.ca/tsmc-65-nm-gp-cmos/>.
- [70] S. M. Sze. *Physics of semiconductor devices*. eng. 2. ed. A Wiley-Interscience publication. New York, N.Y.: Wiley, 1981.

- [71] Robert F. Pierret. *Advanced semiconductor fundamentals*. 2nd ed. Modular series on solid state devices v. 6. Upper Saddle River, N.J: Prentice Hall, 2003. ISBN: 978-0-13-061792-7.
- [72] Mark Lundstrom. *Fundamentals of Nanotransistors*. en. Vol. 06. Lessons from Nanoscience: A Lecture Notes Series. WORLD SCIENTIFIC, Sept. 2017. DOI: [10.1142/9018](https://doi.org/10.1142/9018). URL: <https://www.worldscientific.com/worldscibooks/10.1142/9018>.
- [73] Gernot Paasch and Susanne Scheinert. “Charge carrier density of organics with Gaussian density of states: Analytical approximation for the Gauss–Fermi integral”. en. In: *Journal of Applied Physics* 107.10 (May 2010), p. 104501. ISSN: 0021-8979, 1089-7550. DOI: [10.1063/1.3374475](https://doi.org/10.1063/1.3374475). URL: <http://aip.scitation.org/doi/10.1063/1.3374475>.
- [74] H. Bohuslavskyi et al. “28nm Fully-depleted SOI technology: Cryogenic control electronics for quantum computing”. en. In: *2017 Silicon Nanoelectronics Workshop (SNW)*. Kyoto: IEEE, June 2017, pp. 143–144. ISBN: 978-4-86348-647-8. DOI: [10.23919/SNW.2017.8242338](https://doi.org/10.23919/SNW.2017.8242338). URL: <http://ieeexplore.ieee.org/document/8242338/>.
- [75] Wriddhi Chakraborty et al. “Cryogenic Response of HKMG MOSFETs for Quantum Computing Systems”. en. In: *2019 Device Research Conference (DRC)*. Ann Arbor, MI, USA: IEEE, June 2019, pp. 115–116. DOI: [10.1109/DRC46940.2019.9046346](https://doi.org/10.1109/DRC46940.2019.9046346). URL: <https://ieeexplore.ieee.org/document/9046346/>.
- [76] Robert F. Pierret. *Semiconductor device fundamentals*. Reading, Mass: Addison-Wesley, 1996. ISBN: 978-0-201-54393-3.
- [77] S. Takagi et al. “On the universality of inversion layer mobility in Si MOSFET’s: Part I-effects of substrate impurity concentration”. en. In: *IEEE Transactions on Electron Devices* 41.12 (Dec. 1994), pp. 2357–2362. ISSN: 00189383. DOI: [10.1109/16.337449](https://doi.org/10.1109/16.337449). URL: <http://ieeexplore.ieee.org/document/337449/>.
- [78] S. Takagi et al. “On the universality of inversion layer mobility in Si MOSFET’s: Part II-effects of surface orientation”. en. In: *IEEE Transactions on Electron Devices* 41.12 (Dec. 1994), pp. 2363–2368. ISSN: 00189383. DOI: [10.1109/16.337450](https://doi.org/10.1109/16.337450). URL: <http://ieeexplore.ieee.org/document/337450/>.
- [79] Ali Khakifirooz, Osama M. Nayfeh, and Dimitri Antoniadis. “A Simple Semiempirical Short-Channel MOSFET Current–Voltage Model Continuous Across All Regions of Operation and Employing Only Physical Parameters”. en. In: *IEEE Transactions on Electron Devices* 56.8 (Aug. 2009), pp. 1674–1680. ISSN: 0018-9383. DOI: [10.1109/TED.2009.2024022](https://doi.org/10.1109/TED.2009.2024022). URL: <http://ieeexplore.ieee.org/document/5161294/>.

- [80] Lan Wei, Omar Mysore, and Dimitri Antoniadis. “Virtual-Source-Based Self-Consistent Current and Charge FET Models: From Ballistic to Drift-Diffusion Velocity-Saturation Operation”. en. In: *IEEE Transactions on Electron Devices* 59.5 (May 2012), pp. 1263–1271. ISSN: 0018-9383, 1557-9646. DOI: [10.1109/TED.2012.2186968](https://doi.org/10.1109/TED.2012.2186968). URL: <http://ieeexplore.ieee.org/document/6163383/>.
- [81] Amlan Majumdar and Dimitri A. Antoniadis. “Analysis of Carrier Transport in Short-Channel MOSFETs”. en. In: *IEEE Transactions on Electron Devices* 61.2 (Feb. 2014), pp. 351–358. ISSN: 0018-9383, 1557-9646. DOI: [10.1109/TED.2013.2294380](https://doi.org/10.1109/TED.2013.2294380). URL: <http://ieeexplore.ieee.org/document/6702447/>.
- [82] D.-H. Kim et al. “Extraction of virtual-source injection velocity in sub-100 nm III-V HfETs”. en. In: *2009 IEEE International Electron Devices Meeting (IEDM)*. Baltimore, MD, USA: IEEE, Dec. 2009, pp. 1–4. ISBN: 978-1-4244-5639-0. DOI: [10.1109/IEDM.2009.5424268](https://doi.org/10.1109/IEDM.2009.5424268). URL: <http://ieeexplore.ieee.org/document/5424268/>.
- [83] Shaloo Rakheja et al. “A unified charge-current compact model for ambipolar operation in quasi-ballistic graphene transistors: Experimental verification and circuit-analysis demonstration”. en. In: *2013 IEEE International Electron Devices Meeting*. Washington, DC, USA: IEEE, Dec. 2013, pp. 5.5.1–5.5.4. ISBN: 978-1-4799-2306-9. DOI: [10.1109/IEDM.2013.6724568](https://doi.org/10.1109/IEDM.2013.6724568). URL: <http://ieeexplore.ieee.org/document/6724568/>.
- [84] Shaloo Rakheja et al. “An Ambipolar Virtual-Source-Based Charge-Current Compact Model for Nanoscale Graphene Transistors”. en. In: *IEEE Transactions on Nanotechnology* 13.5 (Sept. 2014), pp. 1005–1013. ISSN: 1536-125X, 1941-0085. DOI: [10.1109/TNANO.2014.2344437](https://doi.org/10.1109/TNANO.2014.2344437). URL: <http://ieeexplore.ieee.org/document/6868269/>.
- [85] C.-S. Lee et al. “A Compact Virtual-Source Model for Carbon Nanotube FETs in the Sub-10-nm Regime—Part I: Intrinsic Elements”. en. In: *IEEE Transactions on Electron Devices* 62.9 (Sept. 2015), pp. 3061–3069. ISSN: 0018-9383, 1557-9646. DOI: [10.1109/TED.2015.2457453](https://doi.org/10.1109/TED.2015.2457453). URL: <http://ieeexplore.ieee.org/document/7174994/>.
- [86] Ujwal Radhakrishna et al. “MIT virtual source GaNFET-RF compact model for GaN HEMTs: From device physics to RF frontend circuit design and validation”. en. In: *2014 IEEE International Electron Devices Meeting*. San Francisco, CA, USA: IEEE, Dec. 2014, pp. 11.6.1–11.6.4. ISBN: 978-1-4799-8001-7. DOI: [10.1109/IEDM.2014.7047034](https://doi.org/10.1109/IEDM.2014.7047034). URL: <http://ieeexplore.ieee.org/document/7047034/>.
- [87] Shaloo Rakheja, Mark S. Lundstrom, and Dimitri A. Antoniadis. “An Improved Virtual-Source-Based Transport Model for Quasi-Ballistic Transistors—Part I: Capturing Effects of Carrier Degeneracy, Drain-Bias Dependence of Gate Capacitance, and Nonlinear Channel-Access Resistance”. en. In: *IEEE Transactions on Electron Devices* 62.9 (Sept. 2015),

- pp. 2786–2793. ISSN: 0018-9383, 1557-9646. DOI: [10.1109/TED.2015.2457781](https://doi.org/10.1109/TED.2015.2457781). URL: <http://ieeexplore.ieee.org/document/7175024/>.
- [88] Shaloo Rakheja and Dimitri Antoniadis. “MVS Nanotransistor Model (Silicon)”. In: (). DOI: [10.4231/D3RR1PN6M](https://doi.org/10.4231/D3RR1PN6M). URL: <http://dx.doi.org/10.4231/D3RR1PN6M>.
- [89] Geoffrey Coram. *Verilog-A: Present Status and Guidelines*. 2013. DOI: <https://nanohub.org/resources/18557>.
- [90] A. Ortiz-Conde et al. “A review of recent MOSFET threshold voltage extraction methods”. en. In: *Microelectronics Reliability* 42.4-5 (Apr. 2002), pp. 583–596. ISSN: 00262714. DOI: [10.1016/S0026-2714\(02\)00027-6](https://doi.org/10.1016/S0026-2714(02)00027-6). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0026271402000276>.

# APPENDICES

# Appendix A

## Analytical Derivation for Fermi Energy Location

$$N_A^- = \frac{N_A}{1 + g_A e^{\frac{E_A - E_F}{k_B T}}} \quad (\text{A.1a})$$

$$p = N_V e^{\frac{E_V - E_F}{k_B T}} \Rightarrow \frac{p}{N_V} = e^{\frac{E_V - E_F}{k_B T}} \quad (\text{A.1b})$$

since  $E_A - E_F = E_A - E_V + E_V - E_F$ , [A.1a](#) becomes:

$$\frac{N_A}{1 + g_A e^{\frac{E_A - E_V}{k_B T}} e^{\frac{E_V - E_F}{k_B T}}} = \frac{N_A}{1 + g_A e^{\frac{E_A - E_V}{k_B T}} \frac{p}{N_V}} = \frac{N_A}{1 + \frac{p}{N_\zeta}} \quad (\text{A.2})$$

where

$$N_\zeta = \frac{N_V}{g_D} e^{\frac{E_A - E_V}{k_B T}} \quad (\text{A.3})$$

Therefore, the charge neutrality equation eq. [4.3](#) becomes:

$$p = \frac{N_A}{1 + \frac{p}{N_\zeta}} \quad (\text{A.4})$$

$$\Rightarrow p^2 + N_\zeta p - N_\zeta N_A \quad (\text{A.5})$$

$$\Rightarrow p = \frac{-N_\zeta \pm \sqrt{N_\zeta^2 + 4N_A N_\zeta}}{2} \quad (\text{A.6})$$

Rejecting the negative solution (because  $p$  is a positive density):

$$\Rightarrow p = \frac{-N_\zeta + \sqrt{N_\zeta^2 + 4N_A N_\zeta}}{2} = \frac{-N_\zeta + N_\zeta \sqrt{1 + 4N_A/N_\zeta}}{2} = \frac{N_\zeta}{2} \left[ \sqrt{1 + \frac{4N_A}{N_\zeta}} - 1 \right] \quad (\text{A.7})$$

Using A.1b:

$$N_V e^{\frac{E_V - E_F}{k_B T}} = \frac{-N_\zeta + \sqrt{N_\zeta^2 + 4N_A N_\zeta}}{2} = \frac{-N_\zeta + N_\zeta \sqrt{1 + 4N_A/N_\zeta}}{2} = \frac{N_\zeta}{2} \left[ \sqrt{1 + \frac{4N_A}{N_\zeta}} - 1 \right] \quad (\text{A.8})$$

$$\Rightarrow E_F = E_V - k_B T \left[ \ln \left( \frac{N_\zeta}{2N_V} \right) + \ln \left( \sqrt{1 + \frac{4N_A}{N_\zeta}} - 1 \right) \right] \quad (\text{A.9})$$

Substituting for A.3:

$$\Rightarrow E_F = E_V - k_B T \left[ \ln \left( \frac{1}{2g_A e^{\frac{E_A - E_V}{k_B T}}} \right) + \ln \left( \sqrt{1 + 4g_A \frac{N_A}{N_V} e^{\frac{E_A - E_V}{k_B T}}} - 1 \right) \right] \quad (\text{A.10})$$

# Appendix B

## Analytical Derivation for Fermi Potential

### B.1 Incomplete Ionization Case:

$$p = n_i e^{\frac{\phi_F}{\phi_T}} \quad (\text{B.1a})$$

$$N_A^- = \frac{N_A}{1 + g_A e^{\frac{\phi_F - \phi_A}{\phi_T}}} = \frac{N_A}{1 + \zeta e^{\frac{\phi_F}{\phi_T}}} \quad (\text{B.1b})$$

where  $\zeta$  is:

$$\zeta = g_A e^{\frac{-\phi_A}{\phi_T}} \quad (\text{B.2})$$

therefore:

$$e^{\frac{\phi_F}{\phi_T}} = \frac{-1 \pm \sqrt{1 + 4\zeta \frac{N_A}{n_i}}}{2\zeta} \quad (\text{B.3})$$

rejecting the negative value (because  $p$  is positive) and multiplying by 1 yields:

$$e^{\frac{\phi_F}{\phi_T}} = \frac{\sqrt{1 + 4\zeta \frac{N_A}{n_i}} - 1}{2\zeta} \times \frac{\sqrt{1 + 4\zeta \frac{N_A}{n_i}} + 1}{\sqrt{1 + 4\zeta \frac{N_A}{n_i}} + 1} = \frac{\left(\sqrt{1 + 4\zeta \frac{N_A}{n_i}}\right)^2 + (1)^2}{2\zeta \sqrt{1 + 4\zeta \frac{N_A}{n_i}} + 1} \quad (\text{B.4a})$$

$$\Rightarrow e^{\frac{\phi_F}{\phi_T}} = \frac{2N_A}{n_i \left( \sqrt{1 + 4\zeta \frac{N_A}{n_i}} + 1 \right)} \quad (\text{B.4b})$$

finally, taking the natural logarithm of both side and substituting for  $\zeta$ :

$$\Rightarrow \phi_F = \phi_T \left[ \ln \left( \frac{N_A}{n_i} \right) - \ln \left( \frac{1 + \sqrt{1 + 4g_A \frac{N_A}{n_i} e^{\frac{-\phi_A}{\phi_T}}}}{2} \right) \right] \quad (\text{B.5})$$

## B.2 Complete Ionization Case:

$$N_A = n_i e^{\frac{\phi_F}{\phi_T}} \quad (\text{B.6a})$$

$$\Rightarrow \phi_F = \phi_T \ln \left( \frac{N_A}{n_i} \right) \quad (\text{B.6b})$$

# Appendix C

## VerilogA Code for CryoMVS Model

```
//  
////////////////////////////////////  
  
//Copyright © 2013 Massachusetts Institute of Technology (MIT)  
  
//The terms under which the software and associated documentation (the  
    Software) is provided are as the following:  
  
//The Software is provided "as is", without warranty of any kind, express  
    or implied, including but not limited to the warranties of  
    merchantability, fitness for a particular purpose and noninfringement  
    . In no event shall the authors or copyright holders be liable for  
    any claim, damages or other liability, whether in an action of  
    contract, tort or otherwise, arising from, out of or in connection  
    with the Software or the use or other dealings in the Software.  
  
//MIT grants, free of charge, to any users the right to modify, copy, and  
    redistribute the Software, both within the user's organization and  
    externally, subject to the following restrictions:  
  
//1. The users agree not to charge for the MIT code itself but may charge  
    for additions, extensions, or support.  
  
//2. In any product based on the Software, the users agree to acknowledge  
    the MIT VS Model Research Group that developed the software. This  
    acknowledgment shall appear in the product documentation.  
  
//3. The users agree to obey all U.S. Government restrictions governing  
    redistribution or export of the software.
```

```
//4. The users agree to reproduce any copyright notice which appears on
    the software on any copy or modification of such made available to
    others.
```

```
//Agreed to by
//Dimitri A. Antoniadis, MIT
//May 27 2013
//
```

```
////////////////////////////////////
```

```
// VerilogA for virtual-source (VS) based self-consistent transport/
    capacitance model for Si MOSFET
// transport model: A. Khakifirooz, et al, p. 1674, T-ED 2009.
// charge model: L. Wei et al, p. 1263, T-ED 2012.
// Implemented on July 15, 2013 by S. Rakheja
// Modified on Sep. 19, 2013 by S. Rakheja
// Modified on July 17, 2015 by S. Rakheja
// List of changes:
// 1. Energy transfer parameter zeta added
// 2. Qinvt definition corrected to use FF instead of FFpre
// 3. DIBL correction normalized
// 4. Ballistic charge calculated for linear potential profile (default)
// 5. Inner fringing charge corrected
// 6. Unused variables removed
```

```
// 1.11 Modified on August 2022 by H. Elgabra
// List of changes: Tested on IDVD and IDVG and it matched 1.10 perfectly
    (a transient test is still needed for Q)
// 1. Edited format to comply with NEEDS requirements
// 2. Edited voltage, or acorss vars, from V(a) - V(b) to V(a,b)
// 3. Made elementary charge 'q' user defined so that it is the same on
    all simulators
//      may have to edit %vt, but so far I don't know where it gets its
    q and kB, low priority anyways
// 4. Removed Tjun in favor of ambient temperature. BTW, the temperature
    is sed on ADE XL in C not K, make the necessary conversion there
// 5. Need to change everything to SI units. this is a time-consuming
    endeavor and will have a URA do it
//      Need to also update CDF to use cadence units
```

```
// 1.14 (Cryo edition) Modified on August 2023 by H. Elgabra
// List of changes:
// 0. Cryo-models added
```

```

// 1. Updates formatting to be more inline nanohub recommended practices
// 2. n -> m
// 3. T in Ff smoothing factor changed to 300
// 4. equations changed to better follow original MVS paper formulation
// 5. mu and vx0 removed and replaced by cryo model
// 6. Vt0 -> VtORT

#include "constants.vams"
#include "disciplines.vams"

module cmvs_si_1_1_0(d, g, s, b);
inout      d, g, s, b;
electrical d, g, s, b; // external nodes
electrical di, si;     // internal nodes

// Original VS parameters
parameter real      version      = 1.10; // development one with
      matlab // MVS model version
parameter integer   fetType      = 1      from [-1 : 1] exclude 0;
      // type of transistor. nFET fetType=1; pFET fetType=-1
parameter real      W            = 1e-6    from (0:inf);
      // Transistor width [m]
parameter real      Lgdr         = 65e-9    from (0:inf);
      // Physical gate length [m]. This is the designed gate length for
      litho printing.
parameter real      dLg          = 13e-9    from (0:inf);
      // Overlap length including both source and drain sides [m]
parameter real      Cg           = 1.6681e-2 from (0:inf);
      // Gate-to-channel areal capacitance at the virtual source [F/m^2]
parameter real      etov         = 1.3e-5    from (0:inf);
      // Equivalent thickness of dielectric at S/D-G overlap [m]
parameter real      delta        = 0.09      from [0:inf);
      // Drain-induced-barrier-lowering (DIBL) [V/V]
parameter real      m0           = 1.35      from [0:inf);
      // Subthreshold swing factor [unit-less] {typically between 1.0
      and 2.0}
parameter real      Rs0          = 130e-6    from (0:inf);
      // Access resistance on s-terminal [Ohms-meter]
parameter real      Rd0          = 130e-6    from (0:inf);
      // Access resistance on d-terminal [Ohms-meter]

      // Generally, Rs0 = Rd0 for symmetric source and drain
parameter real      Ts           = 25        from [0:inf);
      // SS saturation temperature [K]

```

```

parameter real      sm          = 0.05          from [0:10];
    // Smoothing factor for effective temperature calculation [unit-
    less]
parameter real      NO          = 45e4          from [0:inf);
    // inflection magnitude [m^-2]
parameter real      TO          = 40            from [0:inf);
    // inflection temperature (width) [K (eV)]
parameter real      SO          = 8.8372e-3     from [0:inf);
    // subthreshold swing off-set [mV/dec]

parameter real      Cif         = 1e-10         from [0:inf);
    // Inner fringing S or D capacitance [F/m]
parameter real      Cof         = 2e-11         from [0:inf);
    // Outer fringing S or D capacitance [F/m]
parameter real      beta        = 1.8           from (0:inf);
    // Saturation factor. Typ. nFET=1.8, pFET=1.6
parameter real      phib        = 1.2           from [0:inf);
    // ~abs(2*phib)>0 [V]
parameter real      gamma       = 0.0           from [0:inf);
    // Body factor [sqrt(V)]
parameter real      VtORT       = 0.486         from [0:inf);
    // Strong inversion threshold voltage at 300K, abs(VtORT) [V]
parameter real      alpha       = 3.5           from [0:inf);
    // Empirical parameter for threshold voltage shift between strong
    and weak inversion.
parameter real      mc          = 0.2           from [0.01:10];
    // Choose an appropriate value between 0.01 to 10

    // For, values outside of this range, convergence or accuracy of
    results is not guaranteed
parameter integer    CTM_select = 1             from [1:inf);
    // If CTM_select = 1, then classic DD-NVSAT model is used

    // For CTM_select other than 1,blended DD-NVSAT and ballistic
    charge transport model is used
parameter real      CC          = 0             from [0:inf);
    // Fitting parameter to adjust Vg-dependent inner fringe
    capacitances(Not used in this version)
parameter real      md          = 0             from [0:inf);
    // Punch-through factor [1/V]
parameter real      zeta        = 1.0           from (0:1];
    // Energy-transfer factor

parameter real      a1          = 203           from (-inf:inf);
    // for mu(T)

```

```

parameter real      a2          = 2555.4      from (-inf:inf);
    // for mu(T)
parameter real      p1          = 0.193       from (-inf:inf);
    // for mu(T)
parameter real      p2          = -0.398      from (-inf:inf);
    // for mu(T)
parameter real      Tmu         = 103.5       from (0:inf);
    // for mu(T)
parameter real      smmu        = 20.9        from (0:inf);
    // for mu(T)

parameter real      b1          = 0.71        from (-inf:inf);
    // for vx0(T)
parameter real      b2          = 2.73        from (-inf:inf);
    // for vx0(T)
parameter real      q1          = 0.05        from (-inf:inf);
    // for vx0(T)
parameter real      q2          = -0.24       from (-inf:inf);
    // for vx0(T)
parameter real      Tvxo        = 100         from (0:inf);
    // for vx0(T)
parameter real      smvx0       = 30          from (0:inf);
    // for vx0(T)

parameter real      cVth        = 1e-14       from (0:inf);
    // for VT0(T)
parameter real      theta       = 0.774       from (0:inf);
    // for VT0(T)

#define SMALL_VALUE  (1e-10)
#define LARGE_VALUE  (40)
#define kB           (1.380649e-23)          // [J/K] Boltzman constant |
    source: https://docs.scipy.org/doc/scipy/reference/constants.html
#define q            (1.602176634e-19)       // [C] elementary charge   |
    source: https://docs.scipy.org/doc/scipy/reference/constants.html
#define M_SQRTPI     (1.772453850905516)    // sqrt(pi)
#define phiRT        (0.025851999786436)    // phi300
#define EgRT         (1.124519230769231)    // Eg(300) with band gap
    widening
#define halfEgRT      (0.562259615384615)    // Eg(300)/2
#define RTpow1p5      (5196.152422706632)    // 300^1.5

```

```

real Rs, Rd, Vds, Vgs, Vgsraw, Vgdraw, Vbs, Vdsi, Vgsi, Vbsi, dir;
real Leff, me, phiT;
real m, mphiTe, aphiT, Vtpcorr, etapre, FFpre, ab, Vcorr, Vgscorr,
    Vbscorr, VtObs, VtObs0;
real Vt, Vt_0, smFF, dVt, Vr, Vr0;                                // threshold
    voltage variables
real T, Te, TTs, Vthstar, mVgs, etastar, mS0;                    // cryo-S
    variables (M and N shall change to something more robust)

real eVg, FF, eVg0, FF0, Qref, eta, eta0;
real Qinv, Qinv_corr, vx0, Vdsats, Vdsat,Vdratio, Vdbeta, Vdbetabeta,
    Fsat, Id;
real Vgt, psis, Vgta, Vdsatq, Fsatq, x, den;
real qsc, qdc, qi, kq, kq2, kq4, tol, qsb, qdb, qs, qd, Qs, Qd;
real Qb, etai, Qinvi, dQinv, dibl_corr;
real Qinvs, Qinvd, Qsov, Qdov, Vt0x, Vt0y, Fs_arg, Fs, Fd_arg, Fd, FFx,
    FFy, Qsif, Qdif, Qg, a, Cofs, Cofd;
real A, B;

real W_d, Lgdr_d, dLg_d, etov_d, Cg_d, Rs0_d, Rd0_d, NO_d, SO_d, Cif_d,
    Cof_d, vx0_d, mu_d, Vt0_d;
real vtTmp, EgAlpha, EgBeta, Eg0, Eg, gAD, phiFCI, phiFII, phiFCIRT,
    phiFIIRT,dphiFCI, dphiFII, Edoping, dVt0, dVtORT;
real fmu, fvxo;
real etas;

analog begin
    // Input: SI, Model is written with device physics unitrs, mainly cm.
    // This section here handles the conversion. the extra variables are no
    // issue to performance
    T      = $temperature;    // in Kelvin
    phiT    = 'kB*T/'q;      // Thermal voltage, kT/q [V]. did not want to
    use $vt
    aphiT   = alpha * phiT;  // Product of alpha and phiT [used as one
    variable]

    W_d     = W      * 1e2;
    Lgdr_d  = Lgdr   * 1e2;
    dLg_d   = dLg    * 1e2;
    etov_d  = etov   * 1e2;
    Cg_d    = Cg     * 1e-4;

```

```

Rs0_d  = Rs0  * 1e6;
Rd0_d  = Rd0  * 1e6;
N0_d   = N0   * 1e-4;
Cif_d  = Cif  * 1e-2;
Cof_d  = Cof  * 1e-2;

// Band gap widening
EgAlpha = 4.730e-4;
EgBeta  = 636.0;
Eg0     = 1.170;
Eg      = Eg0 - ( (EgAlpha * T**2.0) / (T + EgBeta) );
// Vt0 temp dependence
Edoping = 45e-3; // dopant energy. 45mV for boron or phosphorus.
phiFCI  = phiT * ln(cVth / (T ** 1.5)) + (Eg / 2.0);
phiFCIRT= 'phiRT * ln(cVth / ('RTpow1p5)) + 'halfEgRT;
phiFII  = phiFCI - phiT * ln(1.0 + sqrt(1.0 + 4.0 * gAD * exp(
Edoping/phiT ) * cVth/(T ** 1.5) )) + 'M_LN2;
phiFIIRT= phiFCIRT - 'phiRT * ln(1.0 + sqrt(1.0 + 4.0 * gAD * exp(
Edoping/'phiRT) * cVth/('RTpow1p5) )) + 'M_LN2;
vtTmp   = phiFCI + theta*sqrt(phiFII + phiFCI );
dVtORT  = phiFCIRT + theta*sqrt(phiFIIRT + phiFCIRT);
dVt0    = vtTmp - dVtORT;
Vt0_d   = VtORT + dVt0;

// cryogenic S-related formulas are written later

// Trnsport parameters temperature dependence
fmu     = 1.0 / (1.0 + exp( (T-Tmu )/smmu ));
fvxo    = 1.0 / (1.0 + exp( (T-Tvx0)/smvx0 ));
mu_d    = a1*T**p1*fmu + a2*T**p2*(1.0-fmu);
vx0_d   = b1*T**q1*fvxo + b2*T**q2*(1.0-fvx0);

Edoping = 45e-3; // dopant energy. 45mV for boron or phosphorus.

//Voltage definitions
Vgsraw  = fetType * V(g,si);
Vgdraw  = fetType * V(g,di);
if (Vgsraw >= Vgdraw) begin //NMOS
    Vds  = fetType * V(d,s);
    Vgs  = fetType * V(g,s);
    Vbs  = fetType * V(b,s);
    Vdsi = fetType * V(di,si);
    Vgsi = Vgsraw;
    Vbsi = fetType * V(b,si);
    dir  = 1;

```

```

        gAD = 2.0;                // Donor Degeneracy factor in Si
end else begin                    // PMOS
    Vds = fetType * V(s,d);
    Vgs = fetType * V(g,d);
    Vbs = fetType * V(b,d);
    Vdsi = fetType * V(si,di);
    Vgsi = Vgdraw;
    Vbsi = fetType * V(b,di);
    dir = -1;
    gAD = 4.0;                // Acceptor Degeneracy factor in Si
end

// Parasitic element definition
Rs = Rs0_d * 1e-4 / W_d;                // s-terminal
resistance [ohms]
//Rd = Rd0_d * 1e-4 / W_d;                // d-terminal
resistance [ohms] {Uncomment for asymmetric source and drain
resistance.}
Cofs = (0.345e-12 / etov_d) * dLg_d / 2.0 + Cof_d;    // s-terminal
outer fringing cap [F/cm]
Cofd = (0.345e-12 / etov_d) * dLg_d / 2.0 + Cof_d;    // d-terminal
outer fringing cap [F/cm]
Leff = Lgdr_d - dLg_d;                // Effective
channel length [cm]. After subtracting overlap lengths on s and d
side
me = 9.1e-31 * mc;                // Carrier mass
[Kg]

// Effective temperature (Te) calculation for cryo-T SS saturation
effect
TTs = (T-Ts) / sm / Ts;                // compute exponential factor
for Te

// Uncomment for exact Te                | slower
if(T == 0) begin
    Te = 0;
end else if(TTs >= 100) begin
    Te = T;
end else begin
    Te = Ts * (1.0 + sm * ln(1.0 + exp( TTs )));
end

// Uncomment for approximated Te | faster
//if (TTs >= 5) begin

```

```

//      Te = T;
//end else if (TTs  <= -5) begin
//      Te = Ts;
//end else begin
//      Te = Ts * (1 + sm * ln(1 + exp( TTs )));
//end

//Correct Vgsi and Vbsi
//Vcorr is computed using external Vbs and Vgs but internal Vdsi,
Qinv and Qinv_corr are computed with uncorrected Vgs, Vbs and
corrected Vgs, Vbs respectively.
Vtpcorr = Vt0_d + gamma * (sqrt( abs( phib - Vbs ) ) - sqrt( phib )) -
Vdsi * delta;    // Calculated from extrinsic Vbs

// Fermi-function only used to compute the correction factor
etapre = (Vgs - Vtpcorr) / (aphiT * 1.5);
if (etapre >= 10) begin
    FFpre = 0;
end else if (etapre <= -10) begin
    FFpre = 1;
end else begin
    FFpre = 1.0 / (1.0 + exp( etapre ));
end

ab      = 2.0 * (1.0 - 0.99 * FFpre) * phiT;
Vcorr   = (1.0 + 2.0 * delta) * (ab * 0.5) * exp( -Vdsi / ab );
// Correction to intrinsic Vgs
Vgscorr = Vgsi + Vcorr;
// Intrinsic Vgs corrected (to be used for charge and current
computation)
Vbscorr = Vbsi + Vcorr;
// Intrinsic Vgs corrected (to be used for charge and current
computation)
Vt0bs   = Vt0_d + gamma * (sqrt( abs( phib - Vbscorr ) ) - sqrt( phib
)); // Computed from corrected intrinsic Vbs
Vt0bs0  = Vt0_d + gamma * (sqrt( abs( phib - Vbsi      ) ) - sqrt( phib
)); // Computed from uncorrected intrinsic Vbs
Vt       = Vt0bs - Vdsi*delta;
// Computed from corrected intrinsic Vbs and intrinsic Vds
Vt_0     = Vt0bs0 - Vdsi*delta;
// Computed from uncorrected intrinsic Vbs and intrinsic Vds
smFF     = alpha * 'phiRT;
// Fermi-funciton smoothing factor, aphiT, kept at its 300K
value to perserve smoothness during cryo-T

```

```

dVt      = phiT;
// Range at which Vt changes between strong and weak inversion
FF        = 1.0 / (1.0 + exp( (Vgscorr - (Vt - 0.5*dVt)) / smFF));
FF0       = 1.0 / (1.0 + exp( (Vgsi - (Vt_0 - 0.5*dVt)) / smFF));
Vr        = Vt - FF*dVt;
// Refreence Voltage, a better way to look at it than threshold
voltage: http://dx.doi.org/10.1109/ted.2009.2024022
Vr0       = Vt_0 - FF*dVt;
// Vr but for uncorrected intrinsic Vbs

// S gate-voltage dependence
etastar = (Vgs - Vr) / 0.2 / Vr;
if (etastar <= -6) begin
    Vthstar = Vgs - Vr;
end else if (etastar >= 6) begin
    Vthstar = Vr;
end else begin
    Vthstar = Vr * (1 - 0.2 * ln( 1 + exp( -(Vgs - Vr) / 0.2 / Vr ) )
);
end

// m calculation
mS0      = S0/'M_LN10/phiT; // m due to S0
m        = (m0 + md*Vdsi)*Te/T + mS0;
mVgs     = N0_d * exp( (Vthstar-Vr) / (m * 'kB*T0/'q) );
m        = (m0 + md*Vdsi)*Te/T + mVgs + mS0; // Total subthreshold
swing factor taking punchthrough into account [unit-less],
mphiTe = m * phiT; // Product of m and
phiT [used as one variable]

eta      = (Vgscorr - Vr) / mphiTe; // Compute eta factor
from corrected intrinsic Vgs and intrinsic Vds
eta0     = (Vgsi - Vr0) / mphiTe; // Compute eta0 factor
from uncorrected intrinsic Vgs and internal Vds.
// Using FF instead of
FF0 in eta0 gives smoother capacitances.
Qref     = Cg_d * mphiTe;

//Charge at VS in saturation (Qinv)
etas = -30;
if (eta <= 'LARGE_VALUE) begin
    if(eta <= -28) begin // saturate. why from -28? in PlatGround
        eta = abs(etas)*(1.0+0.01*ln(1.0+exp((eta-etas)/abs(0.01*etas
))))+etas-abs(etas);

```

```

        end
        Qinv_corr = Qref * ln( 1.0 + exp(eta) );
    end else begin
        Qinv_corr = Qref * eta;
    end // possible optimisation here for small value
    if (eta0 <= 'LARGE_VALUE) begin
        if(eta0 <= -28) begin // saturate. why from -28? in PlatGround
            eta0 = abs(etas)*(1.0+0.01*ln(1.0+exp((eta0-etas)/abs(0.01*
etas))))+etas-abs(etas);
        end
        Qinv      = Qref * ln( 1.0 + exp(eta0) );           //
    Compute charge w/ uncorrected intrinsic Vgs for use later on in
    charge partitioning
    end else begin
        Qinv      = Qref * eta0;
    end // possible optimisation here for small value

// Transport equations
vx0      = vx0_d;
Vdsats   = vx0 * Leff / mu_d;
Vdsat    = Vdsats * ( 1.0 - FF ) + phiT * FF;
// Saturation drain voltage for current
Vdratio  = abs( Vdsi / Vdsat);
Vdbeta   = pow( Vdratio, beta);
Vdbetabeta = pow( 1.0 + Vdbeta, 1.0 / beta);
Fsat     = Vdratio / Vdbetabeta;
// Transition function from linear to saturation.

// Fsat = 1 when Vds>>Vdsat; Fsat= Vds when Vds<<Vdsat

// Total drain current
Id = Qinv_corr * vx0 * Fsat * W_d;

// Calculation of intrinsic charge partitioning factors (qs and qd)
Vgt = Qinv / Cg_d;
// Use charge computed from uncorrected intrinsic Vgs

// Approximate solution for psis is weak inversion
if (gamma == 0) begin
    a = 1.0;
    if (eta0 <= 'LARGE_VALUE) begin
        psis = phib + phiT * ( 1.0 + ln( ln( 1.0 + 'SMALL_VALUE + exp
( eta0 ))));
    end else begin

```

```

        psis = phib + phiT * ( 1.0 + ln( eta0 ));
    end // possible optimisation here for small value
end else begin
    if (eta0 <= 'LARGE_VALUE') begin
        psis = phib + (1.0 - gamma) / (1.0 + gamma) * phiT * (1.0 +
ln( ln(1.0 + 'SMALL_VALUE + exp( eta0 ) )));
    end else begin
        psis = phib + (1.0 - gamma) / (1.0 + gamma) * phiT * (1.0 +
ln( eta0 ));
    end // possible optimisation here for small value
        a      = 1.0 + gamma / (2.0 * sqrt( abs( psis - Vbsi ) ));
    end
    Vgta      = Vgt / a;                                // Vdsat in
strong inversion
    Vdsatq    = sqrt( FFO * aphiT * aphiT + Vgta * Vgta); // Vdsat
approx. to extend to weak inversion;
                                                    // The
multiplier of phiT has strong effect on Cgd discontinuity at Vd=0.

    // Modified Fsat for calculation of charge partitioning
    // DD-NVSAT charge
    Fsatq = abs( Vdsi/ Vdsatq )/ ( pow( 1.0 + pow( abs( Vdsi/ Vdsatq ),
beta ), 1.0/ beta ));
    x      = 1.0 - Fsatq;

    // MVS 1.0.1 release uses Tsividis model for implementing DD-NVSAT
charges.
    // In MVS 1.1.0, we use the formulation from L. Wei's 2012 TED paper
referenced in the documentation.
    // Both models are identical.

    // *** From Tsividis (DD-NVSAT)
    // den = 15 * ( 1 + x ) * ( 1 + x );
    // qsc = Qinv * (6 + 12 * x + 8 * x * x + 4 * x * x * x) / den;
    // qdc = Qinv * (4 + 8 * x + 12 * x * x + 6 * x * x * x) / den;
    // qi  = qsc + qdc;
    // Charge in the channel

    // *** From L. Wei (DD-NVSAT)
    A      = pow( (1-x), 2 ) / (12.0 * (1 - (1 - x) / 2.0));
    B      = (5.0 - 2.0 * (1 - x)) / (10.0 - (1 - x) / 2.0);
    qsc = Qinv * (0.5 - (1 - x) / 6.0 + A * (1 - B));
    qdc = Qinv * (0.5 - (1 - x) / 3.0 + A * B);
    qi   = qsc + qdc;

```

```

//QB charge
tol      = ( 'SMALL_VALUE * vx0 / 100.0 ) * ( 'SMALL_VALUE * vx0 /
100.0 ) * me / ( 2 * 'q );
if (Vdsi <= tol) begin
    kq2 = ( 2.0 * 'q / me * Vdsi * zeta ) / ( vx0 * vx0 ) * 10000.0;
    kq4 = kq2 * kq2;

    // Uncomment following lines for linear potential profile
    qsb = Qinv * (0.5 - kq2 / 12.0 + kq4 / 32.0);
    qdb = Qinv * (0.5 - kq2 / 6.0 + 3.0 * kq4 / 32.0);

    // Uncomment following lines for parabolic potential profile
    //qsb      =      Qinv * ( 0.5 - kq2/ 24.0 + kq4/ 80.0 );
    //qdb      =      Qinv * ( 0.5 - 0.125 * kq2 + kq4/ 16.0 );
end else begin
    kq  = sqrt( 2.0 * 'q / me * Vdsi * zeta ) / vx0 * 100.0;
    kq2 = kq * kq;

    // Uncomment following lines for linear potential profile
    qsb = Qinv * (4.0 * (kq2 + 1.0) * sqrt( kq2 + 1.0 ) - (6.0 * kq2
+ 4.0)) / (3.0 * kq2 * kq2);
    qdb = Qinv * (2.0 * (kq2 - 2.0) * sqrt( kq2 + 1.0 ) + 4.0)
        / (3.0 * kq2 * kq2);

    // Uncomment following lines for parabolic potential profile
    // qsb = Qinv * (asinh( kq ) / kq - (sqrt( kq2 + 1.0 ) - 1.0 ) /
kq2);
    // qdb = Qinv * ((sqrt( kq2 + 1.0 ) - 1.0 )
kq2);
end

// Flag for classic or ballistic charge partitioning:
if (CTM_select == 1) begin
    // Ballistic blended with classic DD-NVSAT
    qs = qsc;
    // Calculation of "ballistic" channel charge partitioning
factors, qsb and qdb.
    qd = qdc;
    // Here it is assumed that the potential increases
parabolically from the
end else begin
    // virtual source point, where Qinv_corr is known to Vds-dvd

```

```

at the drain.
    qs = qsc * (1 - Fsatq * Fsatq) + qsb * Fsatq * Fsatq;
    // Hence carrier velocity increases linearly by kq (below)
depending on the
    qd = qdc * (1 - Fsatq * Fsatq) + qdb * Fsatq * Fsatq;
    // efecive ballistic mass of the carriers.
end

// Body charge based on approximate surface potential (psis)
calculation with delta=0 using psis=phib in Qb gives continuous Cgs,
Cgd, Cdd in SI, while Cdd is smooth anyway.
Qb = -fetType * W_d * Leff * (Cg_d * gamma * sqrt( abs( psis - Vbsi )
) + (a - 1.0) / (1.0 * a) * Qinvi * (1.0 - qi));

// DIBL effect on drain charge calculation.
// Calculate dQinv at virtual source due to DIBL only. Then:Correct
the qd factor to reflect this channel charge change due to Vd
// VtObs0 and FF=FF0 causes least discontinuity in Cgs and Cgd but
produces a spike in Cdd at Vds=0 (in weak inversion. But bad in
strong inversion)
etai = (Vgsi - (VtObs0 - FF * aphiT)) / mphiTe;
if (etai <= 'LARGE_VALUE') begin
    Qinvi = Qref * ln( 1.0 + exp( etai ));
end else begin
    Qinvi = Qref * etai;
end
dQinv      = Qinvi - Qinvi;
dibl_corr  = ( 1.0 - FF0 ) * ( 1.0 - Fsatq ) * qi * dQinv / Qinvi;
qd         = qd - dibl_corr;

//Inversion charge partitioning to terminals s and d
Qinvs      = fetType * Leff * ((1 + dir) * qs + (1 - dir) * qd) / 2.0;
Qinvd      = fetType * Leff * ((1 - dir) * qs + (1 + dir) * qd) / 2.0;

//Outer fringing capacitance
Qsov       = Cofs * V(g,si); // possible error, missing * fetType
Qdov       = Cofd * V(g,di);

//Inner fringing capacitance

```

```

Vt0x      = Vt0_d + gamma * (sqrt( abs( phib - fetType * V(b,si) ) )
- sqrt( phib ));
Vt0y      = Vt0_d + gamma * (sqrt( abs( phib - fetType * V(b,di) ) )
- sqrt( phib ));
Fs_arg     = (Vgsraw - (Vt0x - Vdsi * delta * Fsat) + aphiT * 0.5) /
(1.1 * mphiTe);
if (Fs_arg <= 'LARGE_VALUE') begin
    FFx     = Vgsraw - mphiTe * ln( 1.0 + exp( Fs_arg ) );
end else begin
    FFx     = Vgsraw - mphiTe * Fs_arg;
end
Fd_arg     = (Vgdraw - (Vt0y - Vdsi * delta * Fsat) + aphiT * 0.5) /
(1.1 * mphiTe);
if (Fd_arg <= 'LARGE_VALUE') begin
    FFy     = Vgdraw - mphiTe * ln( 1.0 + exp( Fd_arg ) );
end else begin
    FFy     = Vgdraw - mphiTe * Fd_arg;
end
Qsif      = (fetType * Cif_d + CC * Vgsraw) * FFx;
Qdif      = (fetType * Cif_d + CC * Vgdraw) * FFy;

//Partitioned charge
Qs         = -W_d * (Qinvs + Qsovs + Qsif); // s-terminal charge
Qd         = -W_d * (Qinvd + Qdov + Qdif); // d-terminal charge
Qg         = -1 * (Qs + Qd + Qb); // g-terminal charge

//Sub-circuit initialization
I(di,si)   <+ fetType * dir * Id;
I(d,di)    <+ V(d,di) / Rd;
I(si,s)    <+ V(si,s) / Rs;

I(si,b)    <+ ddt( Qs ); // charge term: node si to
node b
I(di,b)    <+ ddt( Qd ); // charge term: node di to
node b
I(g,b)     <+ ddt( Qg ); // charge term: node g to
node b
end
endmodule

```

# Glossary

**cryoT** Cryogenic temperatures. This term refers to the temperature range from absolute  $0K$  to  $\sim 120K$  (or  $-273.15^{\circ}C$  to  $\sim -153^{\circ}C$ ) [3–9](#), [13–17](#), [19–26](#), [31](#), [32](#), [34](#), [40](#), [43–51](#), [54–58](#), [60](#), [62](#), [64](#), [66](#), [77](#), [93](#), [96–98](#), [100](#), [101](#)

**FinFET** A type of Metal-Oxide-Semiconductor Field-Effect Transistor that has a thin vertical fin instead of being completely planar. [9](#), [25](#), [59](#)

**Janis** A Lake Shore pumped  $4He$  cryostat with base a temperature of  $1.5 K$ . The Janis cryostat is shown in Fig. [3.5d](#) [32](#), [33](#)

**quantum supremacy** The point at which a quantum computer has demonstrated its ability to perform a task that is practically impossible or extremely time-consuming for classical computers to accomplish. [1](#), [2](#)

**RT** Room temperature. Typically taken as  $\sim 300K$  (or  $26^{\circ}C$ ) [3](#), [6](#), [8](#), [13–15](#), [18](#), [26](#), [31–34](#), [39](#), [40](#), [44](#), [46–48](#), [50](#), [54](#), [55](#), [58](#), [62](#), [63](#), [66](#), [83](#), [93](#), [96](#), [98](#), [101](#)